CIECA final report to DG TREN of the European Commission concerning contract SUB-B27020B-E3-CIECA-2002-S07.14185 (evaluation of novice driver training schemes in 6 EU Member States, 2002-2004)

EU NovEV PROJECT

Evaluation of post-licence training schemes for novice drivers

FINAL REPORT

October 8, 2004

Project manager: CIECA (International Commission of Driver Testing Authorities

Chief editors: Nick Sanders, Esko Keskinen

NovEV managing organisation in each participating country:

KfV – Austria
GOCA - Belgium
ECF - France
BASt - Germany
ROVG - Netherlands
RACC - Spain
RACE - Spain

TABLE OF CONTENTS

1. SUMMARY OF NovEV PROJECT

Objectives
The objectives of the NovEV project were to:

- Give countries experience in designing and implementing 2nd phase post-licence training for novice drivers
- Evaluate the effects of the training, in terms of changes in the participants' skills, knowledge, attitudes and behaviour. Effects on accident reduction could not be measured due to the small sample sizes and limited duration of the project.

Setting
Pilot projects were set up in Belgium, France, Germany, Netherlands and Spain (2). The new obligatory multiphase driver training programme in Austria was also included.

Design
The training programmes were based on guidelines established in the EU Advanced project. They included group discussions, on-road 'feedback' drives and track modules. The focus of the training was designed to be on the higher levels of driver behaviour; the methods were supposed to be coaching rather than instruction, and participant-centred activities.

Participants
Apart from in Austria, where participation is obligatory, participants in all other countries were novice drivers taking part on a voluntary basis. The exact selection criteria varied from country to country, but the participants were all young drivers (between 18 and 24) with between 4 months and 3 years driving experience. Incentives to participate ranged from an interest in road safety, to a desire to win a prize (e.g. a new car) or to reduce their probationary period by one year (Germany only).

Measurements
Before-and-after evaluation designs with control group were set up in France, Netherlands and Spain (2). Control groups were randomly selected from the original list of participants.
Belgium dropped out of the project. Austria conducted a before-and-after evaluation (no control group) and Germany carried out a process evaluation (single measurement).

Results
Some significant positive changes in self-reported driving behaviour, knowledge and/or risk awareness were found in all countries where a before-and-after evaluation design with control group was used. In one case, this significant positive change was appraised by trainers (on-road audit). A negative trend was found in one training centre in the Netherlands. Participants in the Austrian multiphase were satisfied with the training. The German programme was largely being implemented as planned. There were indications that, in at least two countries, the message of the track training was considered by participants (and seminar leaders and track trainers in Germany) to be skills-based as well as risk awareness-based (unintended effect).

Conclusions
2nd phase programmes can positively influence the behaviour of young drivers. They can also have a negative effect if implemented incorrectly. Programmes on paper can be implemented differently in practice. Proper training-of trainers and ongoing quality control is vital, particularly as far as track modules are concerned.

Recommendations
2nd phase training should address primarily the higher levels of driver behaviour, be participant-centred and spread out over time. The intervention period should ideally be in the first year of independent driving after the licence. More practical examples of useful exercises in class, on-road or track are needed. Training-of-trainers is equally as important. A training programme to develop coaching skills should be developed. Comprehensive, independent quality control is essential to ensure that the goals of the programme are delivered in practice.

2. BACKGROUND

The EU NovEV Project is a successor to the EU Advanced project on post-licence driver and rider training. It's main objectives were to measure the effects of post-licence "2nd phase" training for novice drivers and to give the participating countries experience in implementing such training.

Obligatory 2nd phase training for novice drivers already exists in Finland (1990+), Luxembourg (1996+) and Austria (2003+). Switzerland is expected to join this list at the end of 2005. The purpose of 2nd phase training is to address the 'higher levels' of driver behaviour which are thought to be largely responsible for road accidents, casualties and fatalities in this category of novice (= largely young) drivers. The 2nd phase should also act as a support mechanism during the most dangerous period of driving for novice drivers immediately after the licence. The 2nd phase should, above all, encourage these drivers to evaluate their strengths and weaknesses, analyse their experiences to date and to reflect on the major risks linked to their profile when driving.

The content of 2nd phase training has been largely drawn from theoretical models such as the GDE ('GADGET') matrix which outlines 4 levels of driver behaviour and the elements to be trained on each level (please see annex 1).

NovEV's predecessor, Advanced, laid down guidelines for countries or organisations wishing to introduce 2nd phase training for novice drivers. These guidelines are based on the preceding EU DAN project (description and analysis of novice driver training), experiences of countries where the 2nd phase is obligatory, theoretical models on young drivers and educational techniques, and first-hand experience when visiting a number of post-licence training programmes across Europe. NovEV represented the opportunity to test these guidelines in practice, and to test the assumptions made during the Advanced project. Advanced exposed the complexities of implementing a coherent and effective 2nd phase programme on paper; NovEV was the first opportunity for many countries and organisations to experience these complexities in practice. The countries that rose to this challenge were Austria, with its new, obligatory multiphase programme, Belgium, France, Germany, Netherlands, Spain (RACC club) and Spain (RACE club). All the programmes developed and evaluated in NovEV were short-term pilot projects, with the exception of the Austrian obligatory multiphase and the German voluntary 2nd phase which has been implemented nationwide and will last until the end of 2009.

3. OBJECTIVES OF NovEV AND OF 2ND PHASE TRAINING ACCORDING TO *ADVANCED*

The main objective of NovEV was to test the guidelines set down in Advanced – and sharpened in NovEV – by evaluating the effects of the 2nd phase training on young drivers. The effects were measured in terms of significant changes with regard to the knowledge, skills, attitudes and driving behaviour of the participants in these schemes. In the majority of the countries involved in NovEV, the evaluation was able to measure the changes in the above variables by comparing the situation before the training to after the training. Control groups were used in order to demonstrate that any changes in the test groups were due to the training and not to any other influences during the same time period.

The ultimate proof of the effectiveness of any form of driver training is, of course, to see a reduction in traffic accidents amongst the target group. Due to the small sample sizes (~100 participants) and short time-frame (1-2 years), such a link could not be established and was never an objective of the NovEV project. Rather, any significant sign of change with regard to risk awareness and safety-oriented changes in driving behaviour was to be judged as positive. It is also important to establish that 2nd phase training does not engender counter-productive effects. Experiences in Norway in the 80s showed that accidents amongst the target group increased following the introduction of a 2nd phase in the basic training programme. This showed that certain forms of training, particularly with regard to track-based manoeuvring skills exercises, can produce over-confidence amongst novice drivers. Such a phenomenon was highlighted in Advanced and was stressed throughout NovEV as something to avoid at all costs.

A further objective of NovEV was to give the participating countries invaluable experience in implementing 2nd phase training. This was particularly important because prior experience suggested that such a task should not be underestimated.

The EU Advanced report provided the lion's share of guidelines on 2nd phase driver training for the organisations participating in NovEV. The main body of the Advanced report provided a full analysis of existing post-licence training, highlighting best practice as well as the pitfalls and shortcomings observed before and during the project.

Chapter 10 of the Advanced report focused solely on guidelines for 2nd phase training (please see annex 2). It explains the rationale, goals, principles and golden rules with regard to such training. Following the final submission of the Advanced report, a 10-page checklist was specially designed for the trainers of 2nd phase training (see trainers' checklist in annex 14).

The goals of 2nd phase training, as laid down in the Advanced report, are to:

- *raise awareness of risks* on all 4 levels of driving behaviour (see GDE matrix)
- develop a *sense of self-awareness* amongst participants and the ability to recognise the strengths and weaknesses of oneself and those of other road users
- discuss *the theme of mobility and what it means* for young and novice drivers
- encourage *the group process, i.e.* discussing *driving behaviour* in a social context
- build on / refresh / correct *basic vehicle control skills and driving in traffic*
- help to review and correct misunderstanding of *technical and vehicle dynamic facts*
- develop *new and individual safe driving strategies for the future* (based on the risks identified at all 4 levels of driver behaviour), e.g. safe distances, relationship of driver to passenger, etc.

These goals, combined with a series of general principles and '10 Golden Rules' (see annex 2) were put into practice and tested in the EU NovEV project.

NovEV went one step further by designing a checklist for the designers of 2nd phase training (see training checklist in annex 3). This checklist includes a practical and detailed series of questions related to the design and implementation of the 2nd phase, on issues relating to the programme content, trainers and formalities (manuals, rehearsals, etc) of such training.

In terms of evaluating the 2nd phase, chapter 11 of the Advanced report focused on methodologies for assessing the effects of the training. Further information on evaluation methods was provided to the project partners of NovEV, in the form of Mika Hatakka's doctoral dissertation[1] and The Psychology Research Handbook[2].

[1] Hatakka M. (1998) Novice drivers' risk- and self-evaluations. Use of Questionnaires in Traffic Psychological Research Method Development, General Trends in Four Sample Materials, and Connections with Behaviour. 219p. Annales Universitatis Turkuensis, ser.B, Humaniora. Turku: Painosalama.

[2] Leong & Austin. (1996).The Psychology Research Handbook – a guide for graduate students and research assistants.

4. PROJECT PARTNERS

NovEV began with 7 novice driver training schemes in 6 EU countries. The countries, managing organisations, evaluators and national partners are listed in the table below:

Country	**NovEV Project manager**	**Project evaluator**	**Project partners**
Austria	**KfV** (Austrian Road Safety Board)	KfV	ÖAMTC, ARBÖ (automobile clubs)
Belgium	**GOCA** (driver testing umbrella authority)	GOCA	RACB automobile club, FOD (Federal Government Service)
France	**ECF** (French Driving School)	ECF	MACIF insurance company, French Ministry of Transport
Germany	**BASt** (Federal Highway Research Institute)[3]	BASt	DVR (German Road Safety Council =manager of 2nd phase), German Federation of Driving Instructor Associations
Netherlands	**ROVG** (regional road safety platform of Gelderland)	SWOV (traffic research centre)	CBR, VVCR, ANWB, FAM, BOVAG, NOVEM[4] Traffic Test (research company), Ministry of Transport and Waterways
Spain	**RACC** Automobile Club	INTRAS (University of Valencia)	DGT
Spain	**RACE** Automobile Club	INSIA (Polytechnic University of Madrid)	DGT, Spanish Driving Schools Association, AESLEME (NGO on brain and spinal injuries)

The overall NovEV project was managed by CIECA, the international commission of driver testing authorities. An independent evaluation advisor, Esko Keskinen, from Turku University (Finland) was brought in to assess and to advise the project partners on their evaluation designs for measuring the effects of the training.

[3] In practice, the DVR (German Road Safety Council) is responsible for the voluntary 2nd phase programme in Germany. With regard to the NovEV project, however, it is BASt which has the role of manager.

[4] CBR= Dutch driver testing organisation, VVCR= post-licence training company, ANWB= Dutch automobile club, FAM= Driving schools' Association, BOVAG= Driving schools' Association, NOVEM= Center/Novem

5. OVERVIEW OF NovEV NOVICE DRIVER TRAINING SCHEMES

This chapter of the report precedes the detailed description and analyses of the individual training programmes and evaluations of each participating scheme (chapter 6 and 7). It is designed to provide an overview of each scheme, allowing for a basic comparison of the models and designs used in each country. The guidelines on 2nd phase training (see preceding chapter) allow for some flexibility in terms of the design and implementation of the training, so small differences can be seen in each of the NovEV schemes. Some of these differences have come about for practical reasons, often for evaluation purposes[5].

Table 1 provides a general overview of the training and evaluation designs of each participating scheme. The table provides information on the:

- Training modules
- Length of training
- Selection criteria of participants
- Sample sizes (desired sizes, final training samples, and final evaluation samples)
- Evaluation design
- Control groups
- Data collection methods

Table 2 offers a more detailed overview of the contents of each training module, in addition to the profile of the trainers.

The different training programmes differ above all in terms of length: from 1 day for Spain RACC to up to 5 modules on separate days in the German voluntary 2nd phase. All projects were evaluated according to a before-and-after design with control group, with the exception of Germany and Austria (process evaluation only[6]) and Belgium (see below).

The Belgian scheme:

Due to insufficient numbers of participants in Belgium's 'Cool Driving', and to the lack of an effective evaluation design, a decision was made in January 2004 involving the European Commission, NovEV's independent evaluator and CIECA to cancel the evaluation. This decision was reported in NovEV's 2nd interim report to the European Commission in February 2004. 'Cool Driving' had begun before the NovEV was properly underway; this impeded attempts to ensure the implementation of a proper evaluation design and to give feedback on the content of the training programme, and training of the trainers. Shortcomings were found with regard to all the above aspects, and CIECA considered that there was no alternative other than to remove Cool Driving from NovEV. Some aspects of the Belgian programme were considered positively in the context of the NovEV project; these are reported in the best practice section of the conclusions.

[5] For instance, both Spanish projects accepted novice drivers who had held their driving licence for more than one year, despite the fact that the first year following the licence is the generally accepted intervention period for 2nd phase training. This is because the 2nd phase is considered to be effective only when participants have accrued a certain level of driving experience. In Spain, it is common for drivers to accumulate very low mileage in the first 2-3 years following the licence. As a result, the selection criteria for participants were extended there.

[6] In Germany, this was due to time restraints. In Austria, it is because it is an obligatory measure for all novice drivers there.

Table 1: General overview of training and evaluation designs in participating countries (see annexes 16 and 17 for French and German versions)

Countries	NovEV Project Managers	Evaluators	Training Modules	Length of training	Selection criteria	Desired sample sizes	Number trained	Final sample sizes	Evaluation design	Control group	Data collection methods
Austria	KfV	KfV	Feedback drive – Group discussion/Track training – Feedback drive	All modules to be completed in 12 months	Obligatory measure	NA	NA (obligatory measure)	Process evaluation: 1st feedback drive: 330 Track training & group discussion: 846 Wide scale survey: 991	Process evaluation with single measurement and before and after wide scale survey questionnaires	Novices under old (pre-multiphase) programme	questionnaires
France	ECF/ MACIF	ECF	Group sessions – 2 feedback drives – track training	2 days with 4 month interval	18-23 yrs old, 4-6 months after test	198 for test and control gps	124	124 test gp 87 control gp	Before and after questionnaires (7 mths after 2nd day of training)	From list of participants	Questionnaires Accident data
Germany	BASt	BASt	Group discussion - On-road drive - Group discussion - Track training - Group discussion	5 modules over 8 weeks	Any novice driver within probationary period (min. 6 months experience)	NA	300+ as of end August 2004	70	Process evaluation with single measurement	-	Questionnaires
Netherlands	ROVG	SWOV/ Traffic Test	Feedback drive – Track training – Group discussion – Feedback drive	Single training day	18-25 yr old new drivers, around 6 months after test	200 test gp 100 control gp	99	99 test gp 28 control gp	Before and after questionnaire survey & driving audits (1 mth after training)	From list of participants	Questionnaires driving audits
Spain RACC	RACC	INTRAS	Group discussion – track training – feedback drive (order varied)	Single training day	18-24 yrs old less than 3 yrs driving experience	256 test gp 256 control gp	187	124 test gp 114 control gp	Before and after questionnaire survey (6 mths after training)	From list of participants	questionnaires
Spain RACE	RACE	INSIA	Group discussion – track training – feedback drive	Single training day	1-2 years since test, min. 5000km	198 test gp 198 control gp	77	77 test gp 77 control gp	Before and after questionnaire survey and driving audits (2 audits after training, last at 6 mths after training)	From list of participants	Questionnaires driving audits

Table 2: Detailed overview of content of training modules and profiles of trainers

Countries	Group session: content	Length	Group session: trainer	Feedback drive: content	Length	Feedback drive: trainer	Track training: content	Length (hours)	Track training: trainer	Group sizes per trainer
Austria	Fatal accidents amongst novice drivers Evaluation of individual strengths and weaknesses Adoption of individual strategies for safe driving	2 X 50 minutes	Psychologist	Hazard perception Interaction with other road users Discussion on above using a standard feedback form filled in by trainer	50 minutes driving (X 2)	Driving instructor	1 hour theory: driving dynamics and safety features in cars 5 hours practice: demonstration and experience (seating position, braking, cornering, over and understeering, safety margins, viewing technique)	6	Track trainer (OEAMTC, ARBO, etc)	6-12
France	Presentation: man, vehicle, environment Visual and perceptive illusions Drugs and alcohol: including alcohol simulation test Self-reflection and discussion on training	4 hours spread over 2 days	ECF 'animateur'	Hazard perception Decision-making Discussion on above with trainer and 5 other participants in people carrier.	20 minutes driving per person (X 2)	ECF 'animateur'	Discussion on cars of participants (safety features, maintenance, accessories) Braking distance demonstration Loss-of-control simulator	2	ECF 'animateur'	6
Germany	Exchange of experience Personal strengths and weaknesses Driving context: passengers, distractions, time pressure, etc Alcohol and drugs Discussion on training Personal strategies for safe driving	4,5 hours spread over 3 days	Specially qualified driving instructor	Driving and observation of normal driving style Practising situations already identified as weakness Energy and environmentally-friendly driving Discussion on above with trainer and total of 3 participants	60 minutes per person	Same as group session trainer	Braking exercises (emergency braking, braking distances, slippery surfaces, braking with passengers) Driving around bends (comfortably, with passengers, too fast)	4	DVR accredited track trainer	6-12
Netherlands	Based on video sketches of typical novice driver situations (distractions, peer pressure, multi-tasking, tailgating, etc) Spontaneous discussion on the basis of sketches	1 hour 15 minutes	Track trainer	Driving and feedback from trainer based on normal driving style First drive accompanied by total of 2 participants (second only 1) Discussion between trainer and 1-2 participants on basis of self-assessment and trainer assessment	1 hour per person	Instructor / examiner	ABS/non-ABS experience and braking distances Demonstration of braking distances Driving onto the verge Aquaplaning (demonstration) Driving around bends Safety margins	2	ANWB / VVCR track trainer	8-12
Spain RACC	Presentation on traffic accidents Mistakes and offences: 2 causes of accidents Risk factors for novice drivers Use of passive safety systems (e.g. seatbelts)	1,5 hours	INTRAS seminar leader	Driving and feedback from trainer based on normal driving style Drive accompanied by total of 3 participants Discussion between trainer and 3 participants on basis of trainer assessment and participant-observer assessment	20 minutes driving per person	Driving instructor	Emergency braking exercises (ABS/non-ABS, braking distances, effects of slippery surfaces) Slalom (multi-tasking, driving under pressure, distractions)	1.5	RACC track trainer	6-12
Spain RACE	Accident data Perception of risk Lapses in concentration Speed and its relation to accidents Objects inside the vehicle Alcohol, Drugs and their consequences Security Features Other road users Effect of age, of young people between 18 and 24 years old What to do in case of an accident	4 hours	RACE official / AESLEME rep.	Driving and feedback from trainer based on normal driving style Drive accompanied by total of 3 participants Discussion between trainer and 3 participants on basis of trainer assessment and participant-observer assessment	20 minutes per person	Driving instructor	Seating position Slalom Emergency braking Braking distances / safety margins	2	RACE track trainer / AESLEME rep.	9

6. THE SCHEMES IN DETAIL: TRAINING AND EVALUATION

Chapter 6 features detailed studies of the training and evaluation in each country, in the form of individual reports submitted by each country involved in NovEV. Chapter 7 provides 1-3 page summaries of the reports in this section.

- AUSTRIA

NovEV
Results of the Austrian evaluation

Michael Gatscha, Michael Smuc

1. Summary of project

1.1 Participants

Due to the fact that the multiphase system for novice drivers has been obligatory in Austria since 1st January 2003, a deliberate selection of participants was not necessary. Therefore the group samples for this project were chosen randomly.

1.2 Training programme

The second phase training in Austria consists of the following modules:

- Two on-road feedback drives (before and after the track training)
- A track training (on a closed track)
- A psychological group discussion

1.3 Trainers

The on-road feedback drives for novice drivers is accompanied by driving teachers, the track training is led by instructors and the group discussion is conducted by psychologists.
All involved professions have to fulfil several requirements (e.g. education, age, etc.) which are defined by law.

1.4 Evaluation design and data collection methods

The evaluation design (see Table 1) is based on three levels: a process evaluation for both trainers and participants with regard to the track training and the group discussion, a wide scale survey concerning driving attitudes, beliefs and other self-reported data and statistical data from a file of the Central Licence Register concerning all novice drivers in Austria. The predominant collection method was the usage of questionnaires.

Table 1: Evaluation design and data collection methods

	evaluation type	data collection	when
1a	Process evaluation: participants	questionnaire for MPE (="Multi-Phase-Educated") participants	before & after track training
1b	Process evaluation: trainers	questionnaire for MPE trainers	after track training
2	Wide scale survey	control group (SE="standard" education) from "BASIC"(a previous EU-project): questionnaire intervention group (MPE): questionnaire	before and after the introduction of the multiphase system
3	CLR data	Central Licence Register:	Cut-off date: 1st of April 2004

1.5 Analysing methods

For this evaluation only non-parametric tests were used since basic requirements for parametric test were violated.

1.6 Results and conclusions

Novice drivers who completed at least two modules of the multiphase system were generally satisfied with the whole measurement although it is obligatory. This circumstance can be interpreted as evidence for high acceptance the multiphase system in Austria.

For the track training day, most participants mainly expect to learn to master risky situations better. Also the practical part of the track training day was assessed as most applicable for every

day driving. Furthermore, the results show a different view on the importance of several skills between instructors and participants: For example, the ability to correct a skidding car was rated significantly more relevant for real traffic for novice drivers than for instructors, although all skills were considered as very important for safe driving. Nevertheless, it can be concluded that participants may have received a counterproductive message concerning traffic safety during the track training, i.e. that safe driving is based on manoeuvring skills rather than on an anticipatory driving style.

The results of the wide scale survey show that the reduction of practical and theoretical hours of the standard education didn't have statistical significant influence on the pass-rates (number of attempts) of the driving exam, neither on the theoretical test nor on the practical test.

No big differences were found between standard-educated and multiphase-educated novice drivers concerning self assessment of driving style and driving behaviour, offences or accidents. The only differences occurred regarding female persons: they described themselves as more careful drivers and reported less speeding offences.

2. Introduction & Background

The Austrian multiphase programme started on January 1 on an obligatory basis. As such, the Austrian scheme is not a pilot project as in the case of the other schemes in the NovEV project.

For the multiphase programme to become politically acceptable, the new programme could not cost more than the original 'first-phase' driving test. As a result, the basic minimum training requirement for the theory test was reduced to 32 hours (from 40) and for the practical test to 18 (instead of 20). The costs for the additional phases are as follows: track training: 109 Euro; psychological group discussion: 26,40 Euros; feedback drives: normal price for driving lessons.

There are 3 training modules - during the year following the driving test - in the new multiphase programme (see Description of modules). If the novice driver fails to attend these 3 modules, he will receive an automatic warning from the authorities. Within 4 months, his/her licence will be withdrawn. The purpose of the feedback drives (on-road) is to train more advanced skills than in the driving test: hazard perception, social interaction, observing other road users and discussion of their behaviour.

The post-licence period in Austria includes a probation period of 2 years, during which novice drivers must not commit any serious offences such as speeding, drunk driving, overtaking where not allowed and driving through red lights. If this happens, the driver must undergo a psychological examination and is subject to a one year extension of his/her probation period.

2.1 Description of modules

The aim of the "Multiphase Education" (MPE) in Austria is to build a harmonious continuum in driver training (different steps – but on one and the same ladder). This continuum contains on the one hand already implemented models like the driving school education and driving test, and on the other hand new models, such as feedback driving in traffic, road safety training on track, and further education in the psychological aspects of traffic.

The already implemented modules have been improved by intensifying a number of road safety issues (e.g. driving through tunnels, fitness to drive, etc.). The new models have been developed following experiences in other countries.

A general overview of the multiphase education for category B drivers is presented below:

Already implemented modules:

Driving school education

⇓

Theoretical and practical exam

⇓

Driving licence on probation for 2 years and 0,1 % BAC-limit

⇓

New modules:
1.Feedback in road traffic within 2 to 4 months

(two driving sessions of 50 Min.)

⇓

Road safety training and further psychological education within 3 to 9 months (all-day)

⇓

2. Feedback in road traffic within 6 to 12 months (two driving sessions at 50 Min.)

⇓

Probation period finishes after 2 years, provided no serious offences have been committed

The timeframe within which each individual module must be completed offers enough flexibility for the candidate to choose when to take each part. However, he must respect the overall timeframe and for this reason the sequencing of each module has been laid down by law.

Multiphase-education for Category A (motorcycles)

The multiphase education applies also to novice motorcyclists. However, the NovEV project relates only to category B drivers so no further explanation of the motorcycling education is required here.

Driving license with "L17" training

In Austria, young learner drivers also have the option of taking the L17 model. According to this model, drivers can begin their training already at 16 years old, instead of the usual 18. Training is a carefully structured balance of driving school education and lay instruction (covering a minimum of 3000km driven before the test is taken).

Overview: Multiphase-education for „L17“

Already implemented modules:

Extensive driving education with age 16 to 17 with the sign „L17 trainee drive“

⇓

Theoretical and practical test over the age of 17

⇓

Probationary driving licence until 18

⇓

New modules:

Road safety training and further psychological education within 12 months (all-day). Generally taken between 17 and 18 yrs.
Simultaneously: Removal of the L17 speed-limit and the obligation to show “L17” sign

⇓

Feedback in road traffic within 6 to 12 months (Two driving sessions at 50 min.)

⇓

Over the age of 18: 2 further years of probation time and 0,1 ‰ BAC-limit

⇓

Probation period finishes after 2 years, provided no serious offences have been committed

3. Participants

A deliberate selection of participants was not necessary for Austria within the NovEV project, because the multiphase has been mandatory in Austria since 1st January 2003. Therefore, the samples were chosen at random in terms of selection criteria. For the exact numbers of sample sizes, please refer to Table 5: Sample size before and after cleanup.

Selection bias and social desirability bias

Novice drivers who participated in the process evaluation may have been self-selected, because only participants who did the track training at ARBÖ and ÖAMTC-centres were included for the before-after analysis. Nevertheless, this selection bias should be a small one as the group sizes were relatively big and spread all over Austria, despite leaving out participants who did the track training at centres other than ARBÖ or ÖAMTC.

Another approach to keep the social desirability bias out of the process evaluation is to prevent participants from filling out the questionnaire in the same way before and after the training. Therefore the group was split up, in such a way that one group filled out the questionnaire before the training, and the other group afterwards. In order to reduce the social desirability bias when answering the questions, the questionnaire was handed out to the participants at the very end of the courses.

4. Training Programme

4.1 Detailed description

4.1.1 Feedback drives

The most important topics in the feedback drives are the participants' observation skills and coping with mistakes from other road users in order to achieve a defensive accident-avoiding driving style. This can be achieved by optimizing social interaction with other road users. Therefore a standardized feedback-form for trainers was developed, as a basis for discussion with the trainee.

The first and second feedback drive must contain the following (extract):

- Check of seating position and steering technique
- Economical driving
- Driving through tunnels (if possible)
- Overtaking
- Viewing technique
- Narrative driving for a minimum of 10 minutes
- Hazard avoidance training
- Discussion regarding behaviour in tunnels in emergency scenarios
- Discussion regarding necessity and dangers which distract attention from driving

After the on-road session, the individual strengths and weaknesses of the participant should be discussed in order to give constructive feedback to improve safe behaviour in traffic. For better memory retention, the participant receives a results form. For this project, the standardized feedback form could not be analysed, as the authors didn't receive forms.

4.1.2 Track training

Track training focuses on driving dynamics theory, including 5 hours of demonstration and first-hand experience. Emphasis is placed on defensive driving and avoiding self-overconfidence. The group size for the track training day (which also includes the psychological group discussion) is defined by law: a minimum of 6 participants and a maximum of 12.

Primary Goals:

- Through the training the participant should gain knowledge and skills in order to be a better and safer driver
- A more aware and safer driving behaviour should be attained, especially in real traffic
- The participant should understand the importance of observation skills and a defensive driving style in order to recognize risky situations early

With this training the participants' awareness and skills should be improved in order to anticipate critical situations as early as possible; to avoid them as early as possible and in the right manner, and finally to master them or minimize the consequences of such situations.

Contents of theoretical section (one hour):

- Basic physics of driving
- (Emergency) braking techniques
- Possible vehicle reactions on bends
- Causes of over- and understeering
- In-car safety elements

Contents of the practical part (five hours):

- Proper seating position & driving on slalom course

Goal: Understanding the importance of the right seating position, of wearing a seat belt, and of the right viewing technique

- Braking exercises (emergency braking, braking & swerving, braking on slippery ground)

Goal: Participants explore the meaning of the human factor concerning speed, reaction time and the difficulty of assessing changing (driving) situations.

- Correct cornering & braking on bends

Goal: to show participants the relevance of the right viewing, steering and braking techniques. Participants should also understand the multicausality of car-dependent factors (speed, vehicle type, load, tyre condition, etc.) and how vehicles react on bends.

- Correction of an over- and understeering vehicle

Goal: To realise what factors cause under- and oversteering. At the end they see that the best way to manage difficult situations is to avoid them with an anticipatory driving style.

- Assessment of correct safety margins:

Goal: Participants should estimate what safety margins to keep from vehicles driving in front of them at a speed of 50 km/h. The participant has to drive his car in the parallel lane behind the instructor's car at 50 km/h. Suddenly the instructor carries out an emergency braking manoeuvre. In most cases, the participant following closely behind will not be able to brake in time and will pass the instructor's car. If he had been directly behind the car in front, he would have hit it. Participants will see the importance of keeping the right safety margins.

4.1.3 Psychological group discussion

The group discussion lasts for two hours and confronts participants with accidents and risks. The first part of the discussion is presentation-oriented, where psychologists show the distribution of fatal accidents in the age group of 18-24 (participants' age group). It should be deduced by the participants that the most dangerous factor in traffic is the driver him/herself. He or she is the weak point in the system. Therefore psychologists must work on individual beliefs and opinions to raise self-reflecting questions from participants. In the second part the goal is to motivate youngsters to find answers regarding their own strengths and weaknesses, such as "What type of accident would most likely happen to me?" and what specific responsibility they are ready to take on in traffic.

Examples of common issues:
One topic of the group discussion is set on level 3 of the GADGET-Matrix: What factors are really responsible for driving fast in traffic and what are the main disadvantages of a fast and risky driving style when in a hurry? Here the particpants should consider the relevence of several factors (e.g. acceleration and top speed of car, density of traffic, traffic light switching

from green to red, unexpected incidents such as accidents, traffic jams, police checks, etc.) which play the most influental roles when driving in traffic.

Another important focus of the discussion is set on hazard perception, e.g. "what does a flashing green light at traffic lights mean to novice drivers and what does it mean for experienced drivers?" What is going to happen in the next few seconds? The questions should lead to a discussion where hazard perception, observation skills, distance-keeping and speed adaptation are the main subjects. (N.B. Specific situation in Austria: Before the yellow light is engaged, the green light flashes four times to pre-inform the driver). Another example is to elaborate the optimal approach to a pedestrian crossing.

4.2 Locations
The feedback drives are carried out by every driving school which fulfil the requirements set down by law. Practically speaking, feedback drives take place all over Austria.

The track training must be held in specific training areas, known as driving safety centres ("Fahrsicherheitszentren"). These places (about 25 all over Austria) are maintained by associations of driving schools, the Austrian Automobile, Motorcycle and Touring Club (ÖAMTC) and the Austrian Association of Motorists, Motor-Cyclists and Cyclists (ARBÖ).
For this project, data was provided by ÖAMTC and ARBÖ centres.

4.3 Facilities required
For the track training areas, considerable facilities are required: These requirements (e.g. minimum size of the centre, emergency exit areas, existence of classrooms, skid plates, etc.) are mandatory according to Austrian Law. Therefore, a specific commission was founded in order to define and monitor these basic requirements.
For the psychological discussion, a classroom (for a minimum of 13 people) with the following facilities is required:

- Overhead-projector
- Flipchart
- Optional: Laptop computer and beamer, TV-Set with VCR

5. Trainers

5.1 Selection and training of trainers
Most of the requirements for trainers are defined by law, regardless of which profession:

Psychologists:
Group discussion trainers (certified psychologists) must be course moderators of driver-rehabilitation courses or traffic psychologists according to Austrian law. It is also possible that psychologists who are still in training (and are not course moderators or traffic psychologists yet) lead group discussions, if they attend at least 3 track training sessions beforehand.

Track trainers:
Track trainers have to fulfill the following requirements:

Minimum age of 24 years, at least 5 years driving experience, no offences for the past 5 years (especially traffic-related to alcohol, drugs, excessive speeding and cases of hit-and-run); no non-traffic related sentences of more than 6 months.

Track trainers must complete a theoretical and practical training, depending on course type (single or multi-track vehicles) of 16 hours and they have to attend at least three safety training sessions for each course type.

They receive a further one-day psychological group discussion (8 hours): topics: visual perception, basics of learning and information processing and basics of traffic psychology focusing on novice drivers and their behaviour).

On-Road trainers:
Driving instructors must fulfil at least these requirements (by law):

- Being an experienced driving instructor who works in a driving school with 20 hours of special seminars on self-experience or group dynamics, or
- one year work-experience in a driving school and completion of a 12 hour-special training, or
- to be a driving instructor (other than driving school instructor) with three years of work experience with completion of a 12 hour-special training
- Obligatory attendance of a training session including conversation between themselves and participants is also required. A specific topic addressed is difficult communication, especially if the participant's opinion of his or her driving style is different from that of the driving instructor. The driving instructor should be able to give the right constructive feedback.

5.2 Specific training:

Group discussion trainers (Psychologists):

A voluntary 2-day-training including learning and practising moderation, convincing and transmitting psychological "tools". The contents are focused on the needs and interests of the target group, such as typical accident risks on the one hand, and individual accident risks on the other. Topics include self-overestimation, lack of social responsibility, sensation-seeking tendencies in traffic, etc.

Track trainers:

Additional training is only obligatory if the track trainers have led less than 40 courses of track training. The additional training is a two-day course based on teaching didactics (presentation & self-presentation, moderation and education).

On road trainers (driving teachers): no further training was needed.

6. Experiences implementing the training

To implement the multiphase system (or multiphase system) in Austria, a commission consisting of representatives of the following members has been founded:

- Austrian Association of driving schools,
- Austrian Automobile, Motorcycle and Touring Club (ÖAMTC),
- Austrian Association of Motorists, Motor-Cyclists and Cyclists (ARBÖ),
- Austrian Road Safety Board (KfV), and
- Austrian Ministry of Transport, Innovation and Technology (BMVIT)

The main goal of this group is to define standards for organising and implementing the single modules of the second phase by law. The specific requirements and topics have been identified for every single phase (feedback drives 1 and 2, track training and psychological group discussion) in order to build a harmonious continuum for novice drivers and for all trainers involved (driving instructors, track trainers and psychologists). The requirements for the facilities were defined by law. All demands and specifications for the multiphase system can be

found in the Austrian regulation of the driving license law based on § 13a - d (Führerscheingesetz – Durchführungsverordnung).

Nevertheless it is possible that, with regard to track training, variations between single course providers (ÖAMTC, ARBÖ, Associations of single driving schools) do occur. The psychological role is provided by five institutions all over Austria. As a consequence, the approach to group discussions may also vary between these companies. Therefore a handbook for quality assurance is in progress in order to ensure quality between all providers.

Based on the initial experiences establishing the multiphase system in Austria, several observations regarding the implementation process can be made:

- the longer the implementation process, the more "players" are involved
- the exact timetable and definition of specific work packages is necessary for every institution involved
- it is advisable to lay down a requirement to evaluate the training by law

6.1 Feedback from organisers, trainers & participants

Standardised feedback of novice drivers was collected within the process evaluation (track training day) about the single modules. Participants were asked to answer questions regarding their feelings concerning the course and how they experienced their first feedback drive. The categories for answering is in accordance with the Austrian grading system (from 1=very good to 5=very poor).
The predominant problem which was expected during completion of the modules, is the initial resistance to the entire idea of this system. This is due to the mandatory participation. Looking at the feedback of novice drivers, the following results were noted:

6.2 Standard education and first feedback drive:

The following figure shows the general satisfaction of novice drivers of the standard education (n=136) in driving schools and the first feedback drive (n=330):

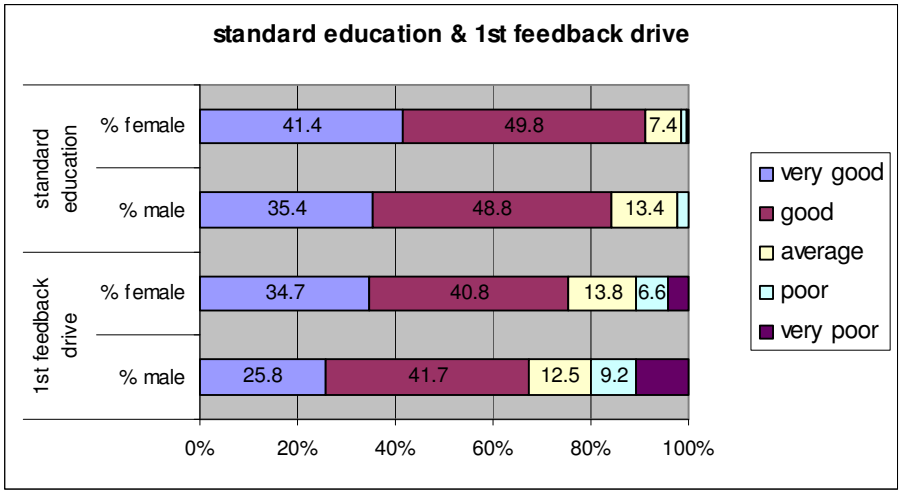


Figure 1: Assessment of standard education and 1st feedback drive

The standard education was better assessed than the first feedback drive (Wilcoxon Test, p=.000).The average value of the standard education was 1.74, whereas the mean for 1st feedback drive was 2.17, which is a satisfactory result. There was also a significant difference

between male and female novice drivers (Mann-Whitney Test, p=.031) concerning the first feedback drive: women assessed this module slightly better than men.

6.3 Track training and group discussion:

The track training was assessed in two separate parts: the theoretical and the practical components. Separately, the psychological group discussion was also assessed. Participants had to evaluate within three categories:

1. The track trainer or psychologist as a person,
2. the presentation/discussion itself and
3. the topics outlined during the training

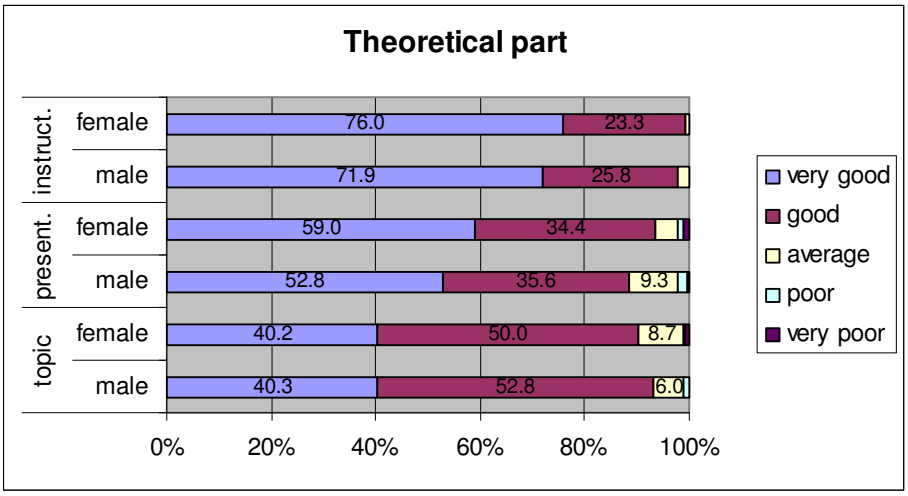


Figure 2: Satisfaction with the theoretical part of the track training

Generally this feedback was pleasing, as most participants answered "very good" or "good", although the distribution of answers is significantly different (Friedman Test, p=.000, n=504) between the three categories: The best grades were given for the instructors, followed by the presentation and the topics.

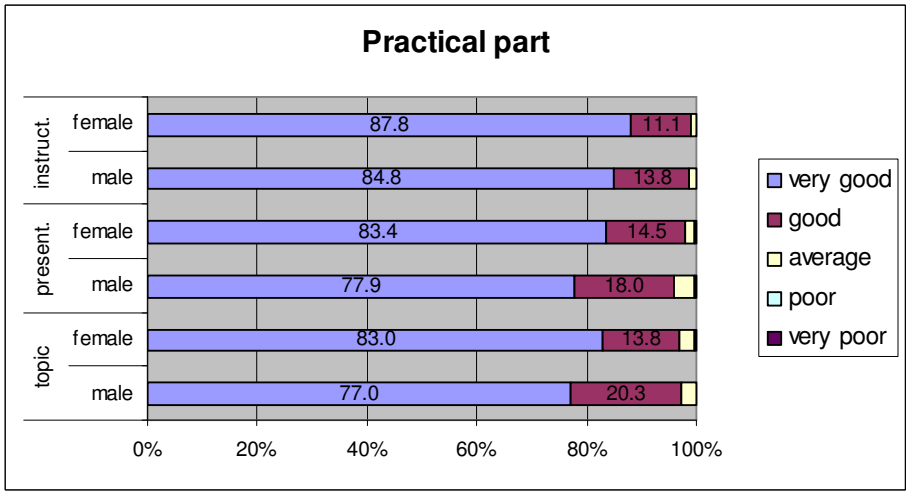


Figure 3: Satisfaction with the practical part of the track training

The practical part of the track training was rated as the best module of the whole multiphase system, as expected. All three categories differ from each other, as seen in the theoretical part (Friedman-test, p=.000, n=506): Best marks were for the instructor, then the presentation, and finally the topics. There were also significant differences between male and female participants concerning the assessment of the presentations (Mann-Whitney test, p=.032) and a trend when assessing the topics (Mann-Whitney test, p=.054) was identified.

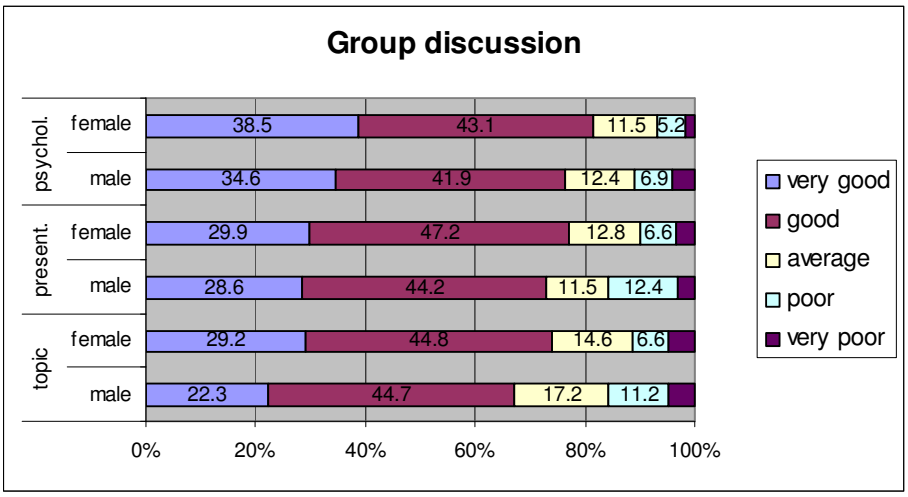


Figure 4: Satisfaction with the psychological group discussion

The group discussion was not as well assessed as the theoretical or the practical part as the figure above shows: The mean value for the psychologist (1.9), for the presentation (2.1) and for the topics (2.2), were significantly different (Friedman-test, p=.000, n=505). This result follows the same order as the theoretical and the practical parts. Differences between male and female clients were found in the "topics"- category (Mann-Whitney test, p=.008).

It can be concluded from the four figures above that the clients were generally satisfied with single modules (feedback drive, track training and psychological group discussion) and gave rather positive feedback (all mean grades between "very good" and "good")although the multiphase is obligatory and can be interpreted as evidence for high subjective acceptance.

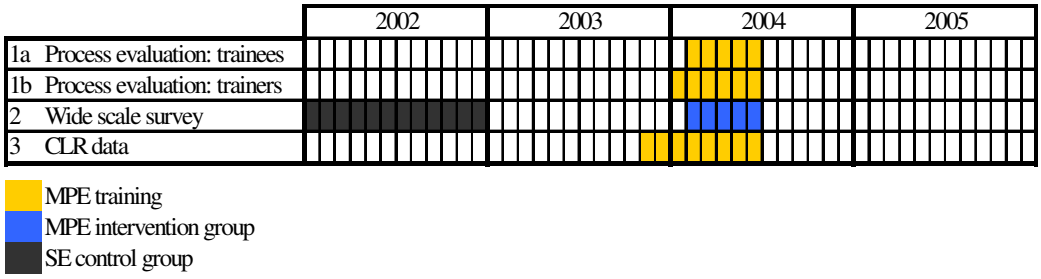
7. Evaluation design and timetables

The evaluation design is based on three levels: a process evaluation for both trainers and participants, a wide scale survey which focuses on the comparison between the standard education system and the multiphase system, and statistical data from all novice drivers in Austria in order to get an impression of the general acceptance. The next two tables show this 3-pronged approach including time frames:

Table 2: Evaluation types, samples description and design (MPE ...multiphase education with standard education; SE ... only standard education).

	evaluation type	source, participants	Design
1a	Process evaluation: participants	ÖAMTC, ARBÖ, driving schools: questionnaire for MPE participants	MPE: before vs. after training session
1b	Process evaluation: trainers	ÖAMTC, ARBÖ, driving schools: questionnaire for MPE trainers	trainers
2	Wide scale survey	control group (SE) from BASIC project: SE intervention group (MPE): sample from CLR data, MPE with at least 2 modules	SE vs. MPE
3	CLR data	Central Licence Register: all novice drivers with at least 2 modules	MPE descriptives

Table 3: Timetable of evaluation



7.1 Process Evaluation

The process evaluation involved questionnaires for both participants and trainers to assess their opinions and attitudes with regard to the training. This part of the evaluation started at the beginning of 2004 and ended in summer 2004.

7.2 Accident statistics

Since the Multiphase Education (MPE) has been obligatory in Austria since January 2003, there is no "synchronized" control group available. If the focus is on accident data, we have to compare two periods: before and after the implementation of the measure. Unfortunately the effects of the measurement have been delayed as the time between the start of the MPE training and identification of the effects has not been long enough. An accident evaluation will be carried out in 2 or 3 years time.

7.3 Wide scale survey

To compare the attitudes of the target group before and after the measure, a two-group design was chosen. A control group (standard education=SE) and intervention group (multiphase education plus standard education=MPE) was selected. The control group consisted of people who took part in the BASIC-project which was carried out in 2002. The intervention group was randomly selected from the Central Licence Register (CLR). Only licence holders who have already completed at least 2 modules of the multiphase education (MPE) were selected. These participants were also selected so that they were "distributed normally" all over the federal states of Austria to fulfil the requirement of representativeness. In principle, the subjects of the

control group and the intervention group were gathered from the same population in terms of age, sex and exposure, in order to be comparable.

In cooperation with the CLR a questionnaire was sent to 2500 novice drivers aged 18 and 19 (only licence holders of category B), and 1142 persons returned the form. This is a response rate of 45.68 %, which far exceeded the expectations. One reason for the high response rate may be that licence holders received a reminder letter, 2 weeks after the first mailshot, to return the form.

The main source for collecting information for this part of the evaluation was, as mentioned before, the use of questionnaires. The questions were concrete and clear and contained one idea per question. The response format was mostly multiple choice except for estimation tasks or self-reported personal data. Another specific point of interest was the development of the pass-rate of licence holders, which was also asked.

7.4 Data file from the CLR

The sample of the Central Licence Register consisted of the whole group of novice drivers in Austria, and the information was provided by the official Department of Statistics in Austria. The data consists of the whole group of novice licence holders for category A and B.

8. Data collection methods

The following table shows all instruments which were used in this study:

Table 4: Overview of data collection types.

	evaluation type	data collection	topics
1a	Process evaluation: participants	questionnaire for MPE participants	Before: Personal demographic data beliefs, self assessment of driving competence, expectations After: Same as before but additional assessment of skills, satisfaction and feedback
1b	Process evaluation: trainers	questionnaire for MPE trainers	Beliefs, assessment of skills
2	Wide scale survey	control group (SE) from BASIC questionnaire intervention group (MPE): questionnaire	Personal demographic data, driving attitudes, beliefs, data about mileage, pass-rates, violations and accidents
3	CLR data	Central Licence Register:	Distribution of MPE-drivers in terms of age, sex and time between issuance of driving licence and completion of single modules

Process evaluation – participants:
The process evaluation was carried out with a before and after track training questionnaire which was filled out either in the morning (before the theoretical part of the track training) or after the group discussion in the evening. Normally the group discussion is the last part of the

track training day although it is possible for the group discussion to take place in the morning. The pre-track training questionnaire for MPE participants included data about: age, sex, driving experience, general, self assessment (socio-demographic data), and also beliefs on fatalism, speeding, distance-keeping, concentration and novice drivers in general. It also contained expectations about experiencing physical boundaries, mastering risky situations and anticipation skills. The post-track training questionnaire was basically the same as the pre-track questionnaire, but additional questions were asked about: anticipating and manoeuvring skills, satisfaction with the course and how practical it was to apply what was learned.

Process evaluation – trainers:
To be able to compare the views of participants and trainers, the instructors also filled out a form. This questionnaire was almost the same as for participants, but shorter. It just consisted of questions on the importance of anticipation and manoeuvring skills for daily traffic, and on the trainers' beliefs. Hence one can observe if there is a gap between the messages delivered by the trainers and the messages received by the participants, or at least the perceived importance of these different skills can be compared across both groups.

Wide scale survey:

The contents of the questionnaire which was used for the wide scale survey were as follows (for used questions see Annex 7):

1. Self-evaluation of individual driving style according to bipolar rating-scale
2. Attitudes towards specific topics
   - speeding,
   - safety margins,
   - risky and competitive driving, etc.
3. Data about driving experience & exposure
   - duration of licence holding,
   - mileage,
   - accidents,
   - offences (alcohol, drugs, speeding, etc.),
   - number of attempts required to pass the driving test.
4. Socio-demographic data
   - age, sex, location, income, job, categories of driving licences
5. Data about personal vehicle
   - type, year of construction, engine power

9. Statistical analysis

9.1 Methods
In feedback-questionnaires the main assumptions of parametric tests are violated: graphical analysis shows that the data is not normally distributed (it is substantially skewed) and tests for the homogeneity of variances show considerable differences.

The low number (of only 4 response categories) is an argument against the existence of an interval scale level in this data, as the availability of fewer categories is less likely to adhere to the properties of the interval scale. The option of using an analogue scale or adding more categories was not chosen in order to avoid the appearance of other problems concerning assumptions of the interval scale.

Due to the fact that the power of a parametric test decreases when its requirements are violated, only non-parametric methods were used for this study. A general weakness of non-parametric

tests is less power when calculating with small sample groups. Fortunately this was not the case in this study and therefore did not apply. As a consequence the following non-parametric tests were used: Mann-Whitney signed rank tests, Kruskal Wallis tests, Wilcoxon tests, Friedman tests and crosstabulations with chi-square statistics and ordinal regression analysis.

The level of significance is reported as trend with $p < .10$, significance with $p < .05$ and high significance with $p < .01$.

9.2 Data cleaning

To ensure a high level of data quality, several steps of data cleaning have been undertaken. The following cases can be distinguished:

- impossible/implausible values: i.e. for exposure (above 150.000 km), age (drivers under 18) and date variables (i.e. mismatch in variables concerning dates: duration of licence holding after date of first module...) [process evaluation:5, wide scale survey before/after: 34/45 cases]
- missing essential grouping variables (mainly gender) [process eval.: 6, survey: 5/4 cases]
- response rates under 60% [process evaluation: 7, survey: 7/3 cases]
- z-values > 2.5 or < -2.5 in self-rating and opinion items (Dollinger & DiLalla, 1996) [process evaluation: 49, survey: 32/22 cases]
- impossible/implausible dependencies between values [process eval.: 0, survey: 73/1 case]

Two rules were applied to deal with impossible or implausible values: If there were indications concerning typing errors in the questionnaire, the corresponding value was changed (e.g. the year of birth "1885" was changed to "1985"). In the case of misstatements the value was deleted. In all other cases, if a noticeable number of values was identified as suspicious, the data concerning this person was discarded. (see overview in Table 5).

Table 5: Sample size before and after cleanup.

	evaluation type	sample size	sample size (after clean up)
1a	Process evaluation: participants	N = 913	N = 846
1b	Process evaluation: trainers	N = 48	N = 48
2	Wide scale survey	NMPE = 1142 NSE = 766	NMPE = 991 NSE = 318
3	CLR data	N = 49336	N = 49297

Concerning the data for the wide scale survey, 297 additional persons of group SE were excluded to provide comparability with group MPE. In order to increase the comparability of these two groups participants holding a licence for more than 1.5 years were excluded.

10. Results

10.1 Process evaluation: participants

A questionnaire was sent to the automobile clubs involved in the track training (ÖAMTC and ARBÖ) to analyse the views and opinions of the participants (licence category A and B) before and after the track training day. In terms of age and sex the groups were equal, the exposure showed a significant difference (chi-square 17.38; df=3; p=.001) in one category (0-1000 km exposure). As a consequence the data was checked to see if a small amount of exposure had an influence or not. It was found that minor exposure had no significant influence on opinions, expectations or attitudes.

10.1.1 General description of the groups:

Table 6: Distribution of age groups on track training day (category B)

	age groups (%)					
	17	18	19	20	21+	n total
male	5.2	72.1	10.1	4.0	8.6	348
female	3.0	64.9	14.5	3.4	14.3	498
both	3.9	67.8	12.6	3.7	11.9	846

A chi-square test revealed a significant difference (chi-square=12.84; df=4; p=.012) of distribution between male and female on the track training day. Male participants were younger. As a result further analyses were done for men and women separately.

Table 7: Driving experience on track training day of the sample (category B)

	exposure (%)				
	0-1000 km	1001-5000 km	5001-10000 km	10001+ km	n total
male	6.8	39.6	29.0	24.6	207
female	21.7	47.0	13.0	18.3	230

As expected a significant difference between men and women (Mann-Whitney test, p=.000) was found. Men did have more driving experience, as the next table also shows:

Table 8: Frequency of usage (category B)

	Since you obtained the driving licence, how often did you drive a car?					
	daily	several times/week	several times/month	rarely	never	n total
male	53.0	35.7	7.8	3.5	0.0	347
female	37.2	39.4	17.1	6.1	0.2	492

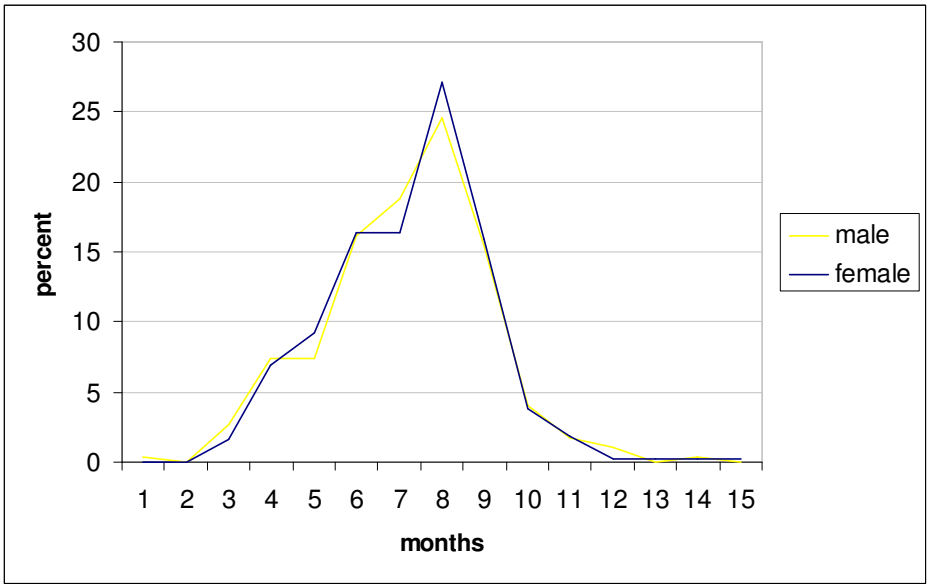


Figure 5: Duration of licence holding (category B) on track training day

Figure 5 shows that the biggest part of novice drivers were on schedule when they attended the track training (most of them hold the licence from 5 to 9 months). This result corroborates the

data of the Central Licence Register mentioned later. There was no obvious difference between sexes.

10.1.2 Description of vehicles

The following table shows that most participants attended the course with their own vehicle, regardless of gender. More female novice drivers than male drivers borrowed a car from the training centre.

Table 9: Description of novice drivers´ used vehicles on track training day

	car for MPE training (%)				
	own car	borrowed	from training center	other	n total
male	60.7	29.8	9.0	0.5	249
female	45.1	37.2	17.0	0.7	212

10.1.3 Expectations for the track training day

Another important question was what participants believe they would experience on the training day. Therefore several questions were asked in the questionnaire before the training started:

"From the track training I expect:

- …to know the limits of my vehicle
- …to master risky situations better
- …to anticipate risky situations faster in order to avoid them
- …to be a better driver compared to drivers who didn't attend the course"

The next figure gives an overview of the results:

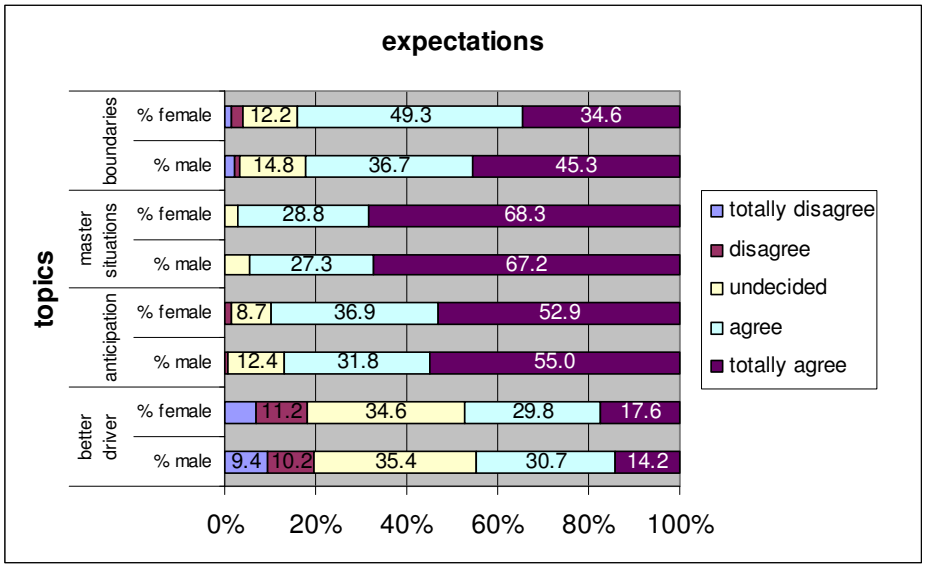


Figure 6: Novice drivers´ expectations of learning possibilities in track training (category B)

Most participants expected to learn to master risky situations better, followed by anticipation skills, knowing the boundaries of ones vehicle. Being a better driver (compared to people who did not attend the training) was least expected. Significant differences between male and female

drivers were not found. All 4 types of expectations were significantly different (Friedman-test, p=.000; n=328) between each other as figure 7 shows:

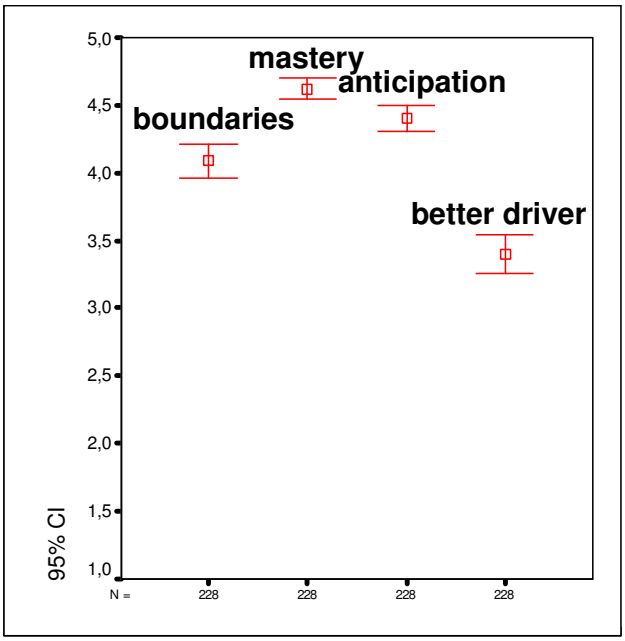


Figure 7: Significantly different expectations of novice drivers concerning learning possibilities concerning track training (category B)

This result gives an indication of how the content of the track training day was delivered by different kinds of media. In general, this circumstance may be an aggravating effect for reaching the goal of the multiphase system, because novice drivers should be prevented from thinking that they can master situations better after the training, which would be an indicator of self-overestimation.

10.1.4 Driving competencies, opinions

Driving competencies

Participants were asked to assess their driving competencies (before and after track training) on three levels by giving marks (1=very good, 5=very poor) according to:

1. vehicle handling (e.g. steering, braking, changing gears, using the clutch),
2. anticipation of risky situations, and
3. an overall assessment

The next three tables show the rating of male and female car drivers before and after the track training and the group discussion:

Table 10: Novice drivers´ self assessment of their vehicle handling skills before and after track training

		vehicle handling (%)						
		very good	good	average	poor	very poor	n total	p
Male	before	22.5	62.8	11.6	3.1	0.0	129	0.083
	after	28.9	61.5	8.7	0.9	0.0	218	
Female	before	10.8	67.5	20.2	1.5	0.0	203	0.245
	after	17.6	60.7	20.3	1.4	0.0	290	

Table 11: Novice drivers´ self assessment of their anticipation skills before and after track training

		anticipation of risky situations (%)						
		very good	good	average	poor	very poor	n total	p
male	before	7.8	68.2	20.9	3.1	0.0	129	0.701
	after	9.3	62.5	27.3	0.9	0.0	216	
female	before	6.0	61.0	29.0	4.0	0.0	200	0.383
	after	5.2	57.1	35.5	2.1	0.0	287	

Table 12: Novice drivers´ overall self assessment of driving competence before and after track training

		overall competence (%)						
		very good	good	average	poor	very poor	n total	p
male	before	5.4	73.6	19.4	1.6	0.0	129	0.364
	after	6.5	66.5	25.6	1.4	0.0	215	
female	before	3.0	53.7	42.3	1.0	0.0	201	0.652
	after	3.6	55.4	39.3	1.8	0.0	280	

More than half of the respondents, males as well as females, rated themselves as having good vehicle handling skills and skills for anticipating risks. Not surprisingly, the same was true for the appraisal of their overall driving competence. The value for the overall assessment was predicted by an ordinal regression model, which revealed that both assessments (vehicle handling and anticipation) had an equal impact on the assessment of the overall competence.

Opinions

With regard to the statement: “A crash depends on fate”, significantly more male novice drivers disagreed with this statement in the group questioned after the training compared to the group questioned before the training. No significant change could be found concerning the female drivers, where the agreement was generally lower in the before group:

Table 13: Opinion before and after track training and psychological group discussion: a crash depends on fate

		"To have a crash is dependent on fate" (%)						
		totally disagree	disagree	undecided	agree	totally agree	n total	p
male	before	19.4	38.0	29.5	10.9	2.3	129	**0.017**
	after	30.3	38.1	21.1	7.8	2.8	218	
female	before	19.4	44.2	27.7	8.3	0.5	206	0.967
	after	24.1	36.4	28.5	9.3	1.7	291	

With regard to the statement: "Driving fast is sometimes more safer", a significant change was detected: After the training the number of females stating that they totally disagree or disagree with this question was higher at the cost of those who were undecided. (Mann Whitney, p=.014):

Table 14: Opinion before and after track training and psychological group discussion: driving fast

		"Driving fast is sometimes safer" (%)						
		totally disagree	disagree	undecided	agree	totally agree	n total	p
male	before	40.3	38.0	18.6	1.6	1.6	129	0.130
	after	45.0	41.7	12.8	0.5	0.0	218	
female	before	44.7	34.5	18.9	1.9	0.0	206	**0.014**
	after	52.2	37.5	8.9	1.0	0.3	291	

About three quarters of the respondents agreed with the view that your concentration is detrimentally affected if you are angry when driving. There were no significant changes, as shown in Table 15:

Table 15: Opinion before and after track training and psychological group discussion: getting angry

		"If you get angry when driving, you´ll drive less concentrated" (%)						
		totally disagree	disagree	undecided	agree	totally agree	n total	p
male	before	0.8	5.4	13.2	41.1	39.5	129	0.212
	after	4.6	7.3	10.6	43.6	33.9	218	
female	before	2.9	4.9	18.5	49.3	24.4	205	0.675
	after	2.8	6.9	17.2	44.8	28.3	290	

Most of the novice drivers disagreed with the statement "with increasing driving experience, one can reduce safety margins…" , while there was no significant change comparing the before and after training groups.

Table 16: Opinion before and after track training and psychological group discussion: reducing safety margins

		"With increasing driving experience one can reduce the safety margin ..." (%)						
		totally disagree	disagree	undecided	agree	totally agree	n total	p
male	before	64.8	28.9	3.9	1.6	0.8	128	0.146
	after	73.4	18.8	5.0	2.3	0.5	218	
female	before	66.5	21.8	8.3	2.9	0.5	206	0.851
	after	68.2	18.2	10.3	2.4	1.0	292	

Both males and females tended to agree with the statement that novice drivers get themselves more often into risky situations. No evidence could be found that drivers questioned before and after the training differ in their level of agreement.

Table 17: Opinion before and after track training and psychological group discussion: Novice drivers and risky situations

		"Novice drivers get themselves more often into risky situations ..." (%)						
		totally disagree	disagree	undecided	agree	totally agree	n total	p
male	before	4.7	14.2	26.8	37.0	17.3	127	0.988
	after	4.6	11.5	30.9	36.4	16.6	217	
female	before	3.9	8.3	42.4	36.1	9.3	205	0.249
	after	5.5	8.9	32.9	40.8	12.0	292	

10.1.5 Ability to apply theory, practical part & group discussion in reality

Participants were asked if they could use something of what was learned in real traffic. The questions were: "For my personal participation in road traffic, I can use something
I learned in the…

- …theoretical part of the training
- …practical part of the training
- …psychological group discussion"

The next figures show the results concerning the sessions on the track training day:

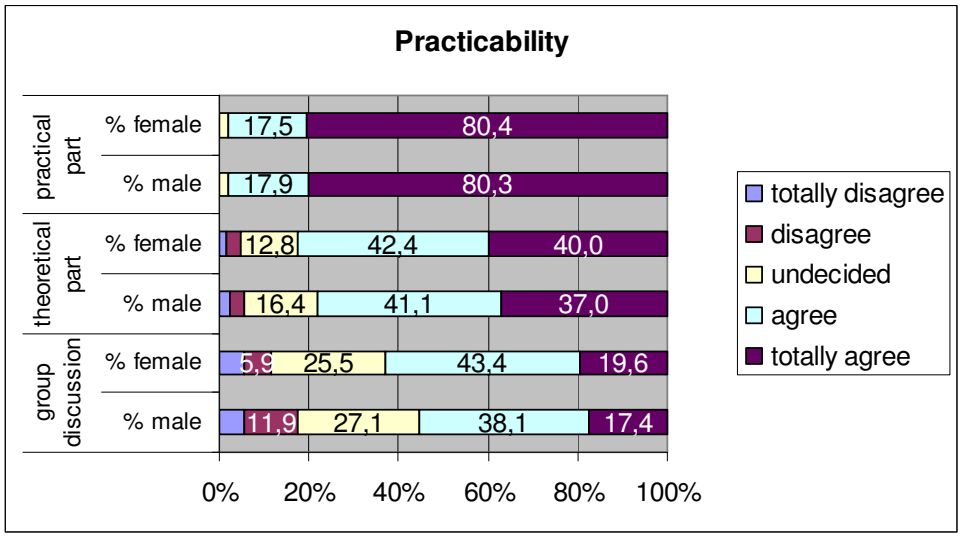


Figure 8: Novice drivers´ assessment of the practical application of single modules on track training day

As the figure shows, all three parts were assessed significantly different (Friedman test, p=.000), as the practical part was judged most applicable, followed by the theoretical part and then the group discussion as also the next figure shows:

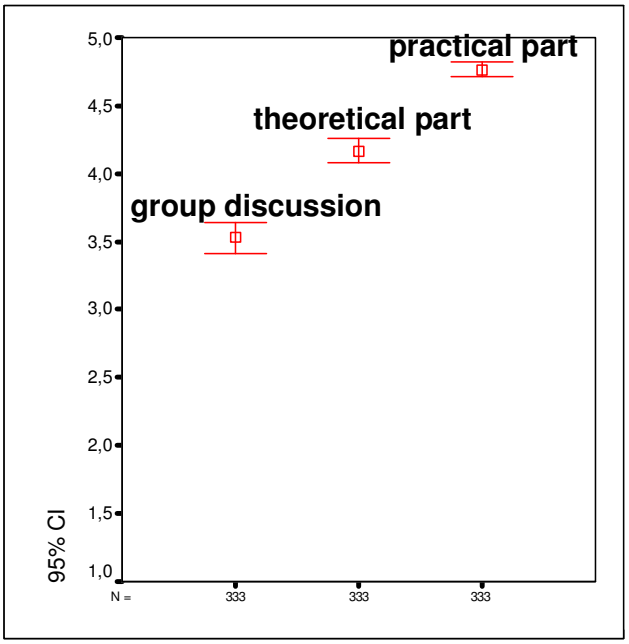


Figure 9: Significantly different assessment of single modules on track training day (category B)

10.2 Process evaluation: trainers

The aim of this analysis was to compare the views of both instructors and participants on the assessment of skills. Therefore participants were asked to rate these skills (5 response categories from 1=not important to 5=very important) by asking this question:

"Assess the importance of the following skills for your personal safety in daily traffic:"

- Recognition of traffic situations which are likely to cause an accident
- Correction of a skidding car
- Avoidance of braking or accelerating in slippery curves
- Selection of speed/distance to avoid a crash in case of emergency braking
- Adaption of speed to situational circumstances
- Correct viewing, steering and braking technique in case of a beginning skid
- Usage of new techniques in real traffic situations
- Avoiding risky situations
- Correct braking and swerving in case of a suddenly appearing obstacle
- Avoidance of driving mistakes which lead to a loss of control over the vehicle

The following results occurred using a Mann Whitney test. The table shows mean values of participants and trainers in ascending (the higher the more important) order:

Table 18: Different assessment between participants and trainers concerning the importance of several skills

Importance of skills	participants		trainers		N total	p
	M	SD	M	SD		
Avoidance of braking or accelerating in slippery curves	4,31	0,77	4,04	0,82	554	**0,022**
Correction of a skidding car	4,47	0,69	3,60	1,07	557	**0,000**
Correct viewing, steering and braking technique in case of a beginning skid	4,55	0,67	4,26	0,99	558	0,088
Usage of new techniques in real traffic situations	4,65	0,58	4,44	0,77	558	0,080
Avoidance of driving mistakes which lead to a loss of control over the vehicle	4,72	0,52	4,98	0,14	559	**0,000**
Avoiding risky situations	4,73	0,55	4,90	0,37	558	**0,032**
Selection of speed/distance to avoid a crash in case of emergency braking	4,76	0,54	4,88	0,33	558	0,227
Recognition of traffic situations which are likely to cause an accident	4,80	0,51	4,94	0,24	558	0,080
Correct braking and swerving in case of a suddenly appearing obstacle	4,83	0,44	4,73	0,49	558	0,078
Adaption of speed to situational circumstances	4,85	0,43	4,94	0,24	559	0,172

The instructors rated some skills as more or less important than participants, although the answers both groups assessed all skills as "rather important" or "very important" as the next figure shows.

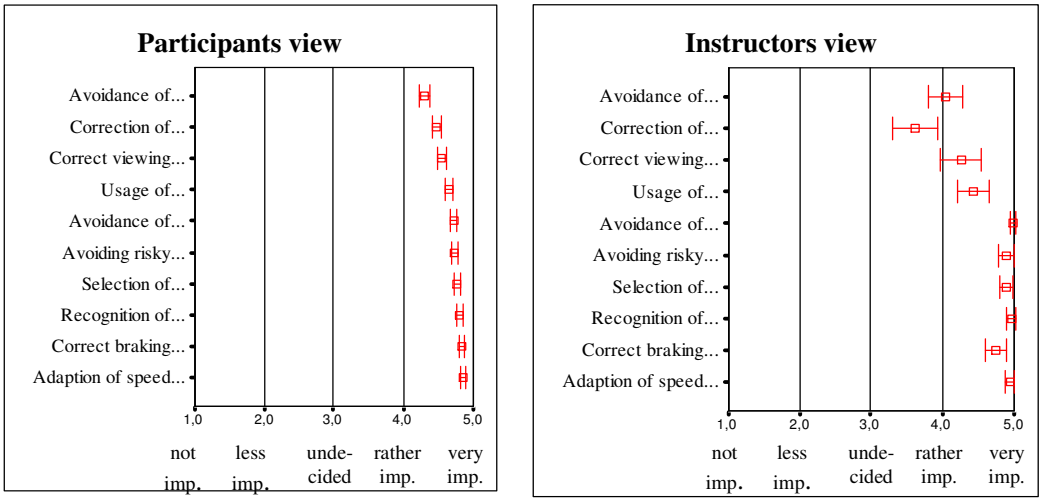


Figure 10: Participants' and instructors' view on skills (after track training)

As figure 10 shows, both instructors and clients rated skills very highly. A factor analysis was carried out in order to identify anticipating and manoeuvring skills, but the reliability of the factors were low due to the considerable skewing of the distribution of the answers.

However, a noticeable difference between instructors and participants was found concerning the importance of correction of a skidding car, whereas instructors rated this skill significantly less relevant for every day driving (Mann Whitney, p=.000). To avoid braking or accelerating on slippery bends seemed more important for participants (Mann Whitney; p=.002), whereas avoiding risky situations (Mann Whitney; p=.003) and driving mistakes which lead to a loss of control over the vehicle (Mann Whitney; p=.000) was a more necessary skill for trainers.

These results led to the conclusion that students may have received the wrong message during the track training, i.e. that safe driving is based on manoeuvring skills, such as correcting a skid, rather than being able to avoid a skid altogether by adopting an anticipatory driving style.

10.3 Wide Scale Survey

As mentioned above, the questionnaire for the wide scale survey was sent during the first part of 2004 to 2500 novice car drivers aged 18 and 19 who already completed a minimum of two modules of the second phase (intervention group; standard and multi-phase education=MPE). The control group (standard education=SE) consisted of a sample of novice drivers (also 18 and 19 years old), which was collected in the year 2002 for another EU-project (BASIC).

10.3.1 General description of the groups

The following sections show the characteristics of the control and intervention group:

Concerning the distribution between gender, no significant differences were found between the control and intervention groups.

Table 19: Distribution of intervention and control groups

		%	n total
male	MPE	41.9	415
	SE	47.2	150
female	MPE	58.1	576
	SE	52.8	168

SE= standard education
MPE= standard education + multiphase (min. 2 modules)

Table 20: Distribution of exposure (intervention and control groups)

		Since passing the driving test, how many kilometres have you driven? (%)									
		< 10000	10000 - 19999	20000 - 29999	30000 - 39999	40000 - 49999	50000 - 59999	60000 - 69999	>= 70000	n total	p
male	MPE	51.3	27.5	13.9	3.2	1.5	1.2	0.2	1.2	411	0.581
	SE	55.2	18.2	14.0	4.2	2.1	1.4	2.8	2.1	143	
female	MPE	78.2	16.8	3.0	0.9	0.2	0.0	0.4	0.6	536	0.601
	SE	71.1	16.4	5.9	3.3	1.3	0.0	0.7	1.3	152	

The table shows that, with regard to the distribution of exposure, the intervention and control group were comparable (Mann Whitney; p=.581, p=.601; uncategorized values).

The distribution over Austria was comparable between groups as no significant differences were observed:

Table 21: Distribution of groups over Austria (intervention and control groups)

		region (%)			n total	p
		east	middle	West		
male	MPE	34.9	47.9	17.2	413	
	SE	39.3	46.7	14.0	150	0.253
female	MPE	36.8	47.3	15.9	573	
	SE	33.7	48.8	17.5	166	0.440

In summary, the subjects of the control group and the intervention group were comparable in terms of sex, exposure and distribution over Austria.

10.3.2 Development of pass-rates

With regard to the driving test, there were concerns that, due to the reduction of practical and theoretical hours (see introduction) of pre-licence driver training, clients would take more attempts to obtain their driving licence. The next two tables show that this apprehension was unjustified:

Table 22: Number of attempts for theoretical test (intervention and control groups)

		How many attempts did it take you to pass the licence category B theoretical test? (%)					
		1x	2x	3x	> 3x	n total	p
male	MPE	91.3	7.5	1.0	0.2	415	0.105
	SE	86.7	12.0	0.7	0.7	150	
female	MPE	91.3	7.8	0.9	0.0	576	0.194
	SE	88.1	9.5	1.8	0.6	168	

Table 23: Number of attempts for practical test intervention and control groups)

		How many attempts did it take you to pass the licence category B practical test? (%)					
		1x	2x	3x	> 3x	n total	p
male	MPE	93.3	5.8	0.5	0.5	415	0.994
	SE	93.3	4.7	2.0	0.0	150	
female	MPE	87.0	12.5	0.5	0.0	576	0.463
	SE	89.3	8.9	1.8	0.0	168	

It can be concluded that the implementation of the multiphase system had no significant statistical influence on attempts to obtain the driving licence.

10.3.3 Timing of modules & other licence categories

Another topic of interest was the level of satisfaction with the timing of each multi-phase module. Therefore clients were asked if the modules should take place earlier, later or if the timing is considered to be satisfactory:

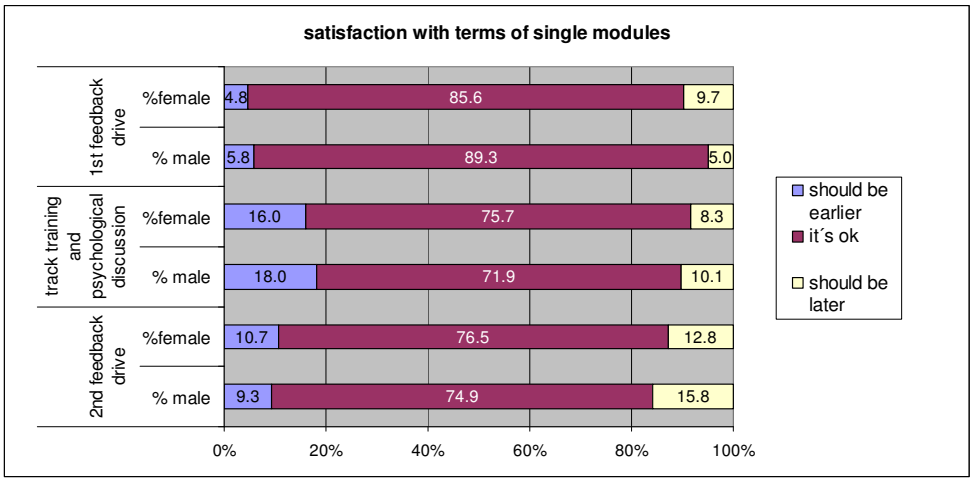


Figure 11: Satisfaction with terms of modules (1st feedback drive, track training & psychological group discussion, 2nd feedback drive)

Most of the respondents, males and females alike, were content with the legal time-frame to complete the courses: Nearly nine out of ten novice drivers did not want to have the first feedback drive module earlier or later. Similarly, the majority of them (about three quarters) supported the existing timeframe for the other modules.

10.3.4 Self rating of behavioural aspects

Clients were asked to assess their own driving style by rating between several aspects of behaviour. From a total of 7 opposite pairs, 2 factors were identified by factor and reliability analysis, as table 24 and 25 show:

Table 24: Factor analysis of opposite pairs of six behavioural aspects

self-assessment of driving style	**factor routine**		**factor care**	
	female	male	female	male
secure/not secure	0,82	0,76	0,01	-0,04
experienced/not experienced	0,78	0,77	0,10	0,10
skilled/not skilled	0,70	0,77	-0,06	-0,13
risky/careful	0,00	-0,03	0,76	0,74
aggressive/non-aggressive	0,04	0,09	0,72	0,74
fast/slow	0,30	0,17	0,61	0,70

Table 25: Reliability analysis of two behavioural factors by gender

reliability analysis	**Cronbach´s alpha values**	
	factor "routine"	factor "care"
male	0.71	0.63
female	0.73	0.59

The first factor “routine” could be described as a safe, experienced and skilled style of driving.
The second dimension found (factor 2=“care”) is related to a careful, non-aggressive and slow driving behaviour.
The factor analysis was carried out for male and female participants, in which an equal structure of factors was identified. Hence a comparison between control and intervention group on the basis of factor scores was calculated. The analysis showed a significant difference regarding to factor “care”: especially female novice drivers described themselves as more careful drivers (Mann Whitney, p=.035).

10.3.5 Opinions

Several safety-related opinions were compared between standard-educated and multiphase-educated novice drivers. The results are shown in the next 2 tables:

Table 26: Comparison of several opinions between intervention and control groups

		Driving at high speed is fun. (%)					
		totally agree	partly agree	partly disagree	totally disagree	n total	p
male	MPE	9.2	56.6	29.2	5.1	415	0.991
	SE	10.7	54.7	28.0	6.7	150	
female	MPE	4.5	50.6	33.2	11.7	575	0.614
	SE	4.8	48.5	32.7	13.9	165	
		I can take more risks when I'm in a hurry. (%)				n total	p
male	MPE	10.4	43.7	34.8	11.1	414	0.319
	SE	6.7	44.0	37.3	12.0	150	
female	MPE	9.7	43.1	32.0	15.1	575	0.315
	SE	8.5	40.0	34.5	17.0	165	
		It is no problem for me to regain control of a skidding car. (%)				n total	p
male	MPE	8.1	58.2	26.9	6.8	409	0.998
	SE	11.7	51.7	29.7	6.9	145	
female	MPE	4.2	45.2	36.2	14.4	569	0.056
	SE	3.8	37.3	39.9	19.0	158	
		Speed limits are mostly unnecessary. (%)				n total	p
male	MPE	2.9	21.2	44.6	31.3	415	0.429
	SE	2.0	18.2	46.6	33.1	148	
female	MPE	1.0	16.1	35.4	47.4	576	0.413
	SE	0.6	15.0	33.5	50.9	167	
		I like to compete in traffic. (%)				n total	p
male	MPE	0.0	6.3	23.8	69.9	412	0.331
	SE	1.3	3.3	21.3	74.0	150	
female	MPE	0.3	2.4	8.6	88.7	573	0.878
	SE	0.0	1.2	10.8	88.0	167	
		Poor condition of tyres does not significantly impair safety. (%)				n total	p
male	MPE	2.7	1.7	10.6	85.1	415	0.358
	SE	7.3	2.7	7.3	82.7	150	
female	MPE	4.7	2.1	11.1	82.1	570	**0.021**
	SE	6.6	5.4	13.8	74.3	167	

The table above shows one significant difference as female 2nd-phase participants agreed more that a bad tyre condition impairs safety.

Table 27: Comparison of several opinions between intervention and control groups (continued)

		When I get more driving experience I may shorten the safety margin to the vehicle in front of me. (%)					
		totally agree	partly agree	partly disagree	totally disagree	n total	p
male	MPE	0.2	2.9	18.7	78.2	412	0.119
	SE	2.0	2.7	23.3	72.0	150	
female	MPE	1.2	3.7	13.6	81.5	574	0.076
	SE	1.2	7.1	16.1	75.6	168	
		Sometimes, I have to terminate overtaking because I have incorrectly estimated the distance involved.. (%)				n total	p
male	MPE	6.5	14.5	32.6	46.4	414	0.589
	SE	6.0	12.7	40.0	41.3	150	
female	MPE	6.1	11.8	30.2	51.9	576	0.227
	SE	6.5	16.1	29.8	47.6	168	
		I always reduce speed in front of school as children might cross the street unexpectedly. (%)				n total	p
male	MPE	68.1	24.4	6.5	1.0	414	**0.050**
	SE	76.0	20.7	3.3	0.0	150	
female	MPE	75.6	20.4	3.5	0.5	574	0.213
	SE	79.8	19.0	0.6	0.6	168	
		When driving without a safety belt I miss something. (%)				n total	p
male	MPE	81.4	15.0	2.4	1.2	414	0.759
	SE	80.7	13.3	4.0	2.0	150	
female	MPE	93.3	5.6	0.4	0.7	571	**0.006**
	SE	86.9	8.9	2.4	1.8	168	
		It is right to withdraw the licence for speeding more than 50 km/h above the limit on rural roads. (%)				n total	p
male	MPE	42.7	34.2	13.8	9.2	412	0.322
	SE	47.0	32.9	12.8	7.4	149	
female	MPE	50.4	30.3	13.3	6.0	571	0.330
	SE	47.6	29.8	12.5	10.1	168	

Another significant difference was noted for seat belt wearing (only for female novice drivers) as mentioned the table.

Generally no major differences between standard-educated novice drivers and multiphase-educated drivers were found. If differences were observed, they were related mostly to female novice drivers. One significant difference was observed: “standard-educated” male novice drivers agreed stronger that they would reduce their driving speed when passing schools. This result indicates that this topic should be more stressed in future courses of the multiphase training.

10.3.6 Offences & Accidents

Offences

Participants of the survey were asked about their traffic offences in order to see if there was a difference between standard and multiphase-educated novice drivers. Offences were categorised in this order:

Offences concerning…

- drink driving up to 0,25 mg/l breath AC
- drink driving > 0,25 mg/l breath AC
- drug impairment
- speeding
- other safety relevant offences
- other (not safety relevant) offences

No statistically significant differences were found, except for speeding:

Table 28: Comparison of self-reported speeding offences between intervention and control groups

		For which offence have you been punished so far an how often ... speeding (%)						n total	p
		0	1	2	3	4	5		
male	MPE	83.6	13.4	2.2	0.2	0.5	0.0	407	0.449
	SE	81.2	13.0	3.6	0.7	0.0	1.4	146	
female	MPE	94.9	4.5	0.6	0.0	0.0	0.0	548	**0.042**
	SE	90.5	5.7	3.2	0.6	0.0	0.0	165	

Female standard-educated novice drivers stated speeding offences more often (Mann Whitney, p=.042). This result may lead to the conclusion that female novice drivers in particular benefit from the multiphase education. Due to the fact that the data of the control group (SE) was gathered in the year 2000, it should be also considered that the development of enforcement measures in Austria may had an effect on the number of detected offences. This effect could not have been estimated since the authors had no data regarding enforcement activities.

Accidents

A total number of 1309 novice drivers gave information regarding their involvement in accidents. For this analysis, all accidents (material and human damage-related accidents) were taken into account:

Table 29: Comparison self-reported accidents between intervention and control groups

		accidents total (%)				n total	p
		0	1	2	3		
male	MPE	76.4	21.2	2.2	0.2	415	0.392
	SE	72.7	25.3	2.0	0.0	150	
female	MPE	79.7	17.7	2.1	0.5	576	0.837
	SE	79.2	17.3	1.8	1.8	168	

Again, no statistical differences were found concerning traffic accidents, either for female or male novice drivers. The next figures show a more detailed description of accidents, despite the fact that there were no significant differences found:

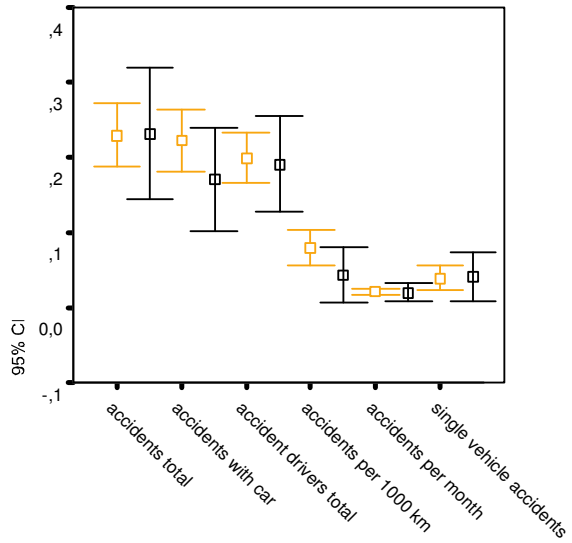


Figure 12: Accidents – female drivers involved (brighter line=SE, darker line=MPE)

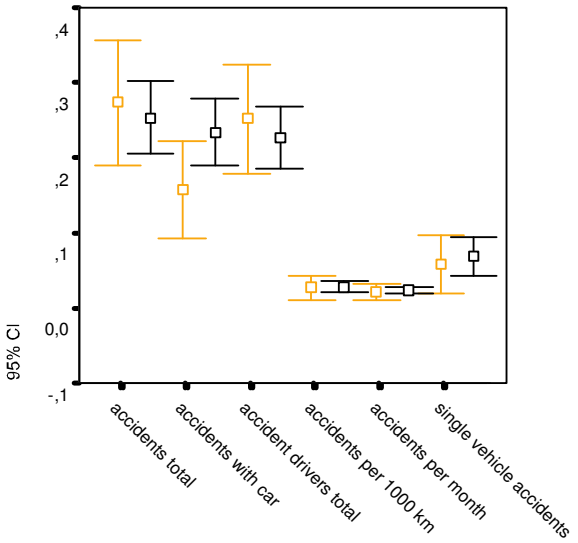


Figure 13: Accidents – male drivers involved (brighter line=SE, darker line=MPE)

No significant differences were found, regardless of which type of accidents is analysed: total number of accidents (with and without car), only accidents with car (material and human damage), accident drivers (drivers who had at least on accident), accidents per 1000 km or month (car only) or single vehicle accidents (motorcycle and car).

10.4 Data file of the Central Licensing Register

To gain an overview of the development of the multiphase in Austria, data was provided from the Central Licence Register (CRL) file with 1st April as a cut-off date. This file contained information about all novice drivers who have to complete the multiphase system, as it has been obligatory since 1st January 2003. The file consisted of the following information:

- Age (partly in categories)
- Sex
- Date of issuance of driving licence
- Date of completion of each module
- Category of driving licence (A or B)

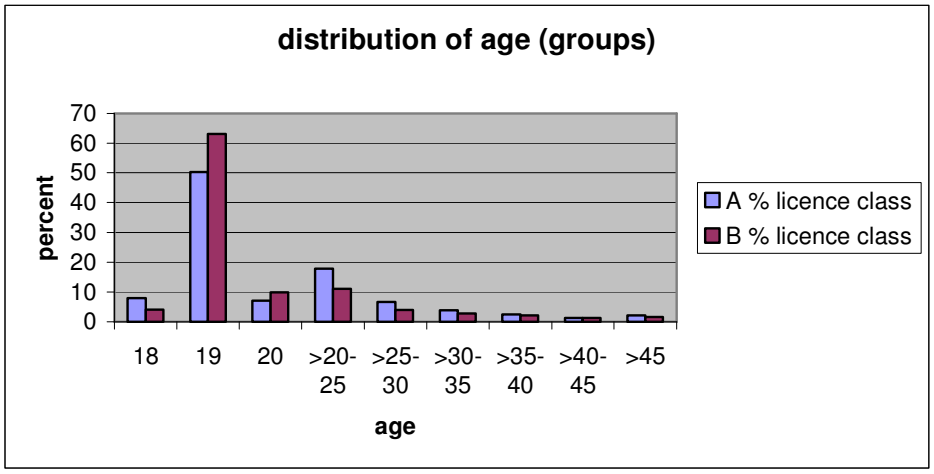
Table 30: Overview of CLR data – distribution of gender and licence category

			sex		
			female	male	total
licence category	A	number	5280	8289	13569
		%	38.91	61.09	100
	B	number	24570	24650	49220
		%	49.92	50.08	100
total		number	29850	32939	62789
		%	47.54	52.46	100

This file contained data from 62.789 people. Men were overrepresented (61%) for category A, whereas for category B the distribution of male and female drivers was nearly equal.

The next table shows the distribution of age depending on the licence category:

Figure 14: Distribution of age and licence category



People who obtained a driving licence for motorcycles (category A) were significantly older. To compare these two distributions a Pearson chi-square test was used. Significant differences (chi-square=1306.5; df=8; p=.000) were observed in every age category, except the age category > 40 to 45 years.

Completion of modules for category B:

As expected the rate of all novice car drivers (n=49297) who completed the first module was much higher (approximately 60 %) compared to motorcycle drivers:

Table 31: Time between issuance and modules for car drivers (not L17!)

			Time between issuance and modules	
Module 1 (1st feedback-drive):	**Number**	**Percent**	**Mean**	**SD**
Completed	29206	59.2	3.4 months	1.2 months
Not completed	20091	40.8		
Module 2 (track training and group discussion):				
Completed	11070	22.5	6.3 months	2.0 months
Not completed	38227	77.5		
Module 3 (2nd feedback drive):				
Completed	2978	6.0	8.6 months	1.8 months
Not completed	46319	94.0		

Many people "chose" the old form of education, probably to avoid the costs of the multiphase education and most novice "multiphase-drivers" started their education later. Apparently this was the main reason why only 6 percent of all novice "multiphase-drivers" since 1st of January 2003 completed all modules by 1st of April 2004. Nevertheless most people completed the modules on schedule.

11. References

Dollinger & DiLalla, 1996. *Cleaning Up Data and Running Preliminary Analyses.* in The Psychology Research Handbook - A Guide for Graduate Students and Research Assistants. Edited by Frederick T. Leong & James T. Austin, University of Tennessee.

- FRANCE

NovEV
Results of the French second phase

Gwenaelle Burguin, Christophe Griffon, Laurent Perrier

1. Summary of project

The participants
396 young members of MACIF insurance company, aged between 18 and 2 years old and having between 4-6 months driving experience, participated in the NovEV project. These young drivers were split into 3 groups : 124 in the experimental group, 87 in the control group and 124 in control group 2. Control group 2 was unaware that it was being monitored, whereas the other two groups had expressed an interest in participating actively in a road safety training programme.

The training
Experience gained in the past by ECF suggested that the programme should be spread over 2 days. These two training days were separated by a 4 month interval. This allowed for more intensive debates and exchanges between the participants.
The training programme contained information, and discussion on different risks (either subjective or objective). It alternated between workshops, on-road sessions and track-based modules. The programme takes into account the hierarchical model of driving behaviour and is particularly focused on levels 3 and 4 of the GDE (goals for driver education) matrix.

The trainers
The whole programme depended heavily on the quality of the discussion and on the pedagogical quality of the training. The 5 trainers used were road safety professionals who were qualified and experienced in giving training to groups of young drivers. They trained in pairs during the entire programme. In order to help them and to retain a coherent approach amongst the different trainers, a trainers' guide was developed especially for this programme.

Feedback on the training
The organisers, trainers and participants all rated the experience positively.

The evaluations
The main objective was to measure and to compare changes in skills, attitudes, knowledge and driving behaviour amongst the participants who actually took the training, and those who did not.
The participants were monitored over a period of 11 months using specially designed questionnaires, as follows :

- Pre-training questionnaire (experimental and control group 1)
- Post-training questionnaire (experimental and control group 1)
- MACIF accident monitoring (for the 3 groups)

Results
Positive changes in the experimental group :
Significant positive change in awareness of risks linked to driving habits (MALES)
Significant positive change in driving skills for defensive driving (MALES)
(Slight) trend towards less frequent risky driving situations (MALES)

Stability of control group.

Conclusion
We can reasonably conclude that the development of the two groups shows an increase in risk awareness in the experimental group. This helps to delay the phenomenon of overconfidence which is so often observed amongst novice drivers.
Otherwise, the control group, which was followed statistically but not involved in the training, remained stable in its results, despite its clear investment in road safety (by wanting to take part)

FOREWORD

The **ECF (Ecole de Conduite Française = French Driving School)** is the leading road safety education organisation in France and has been organising research into road users' education since 1993 with the MACIF (France's biggest car insurance company).

These initiatives focused particularly on post-licence activities and young drivers. Due to this experience, the French **Ministry of Transport** requested ECF and MACIF to participate in the EU NovEV project.

The main partners in this initiative are thus ECF, MACIF and the Ministry of Transport.

2. Participants

2.1 Selection

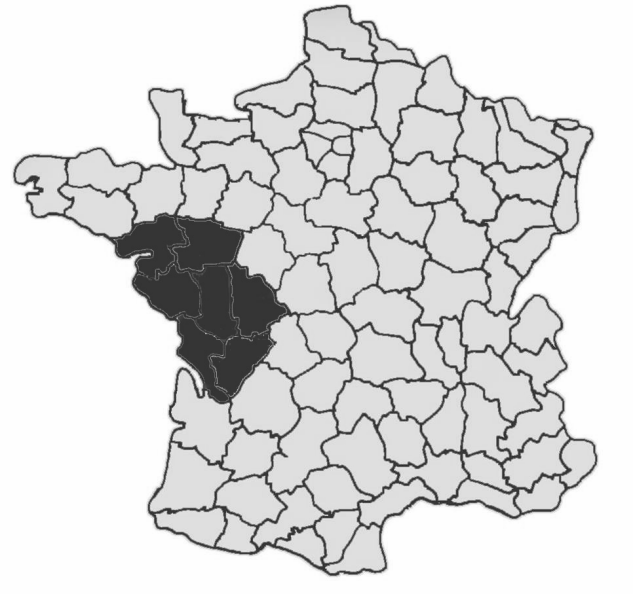
The participants are all members or children of members of the MACIF insurance company. This initiative was a voluntary one designed for young drivers.

More than a thousand young drivers aged between 18 and 23 years old were contacted by letter regarding the project.

It was considered important for the young drivers to have already had a certain amount of driving experience before following the post-licence training in question.

4 to 6 months after passing the driving test, the young drivers were contacted by ECF. This period is the one in which young drivers gain in confidence and also increase their risk-taking.

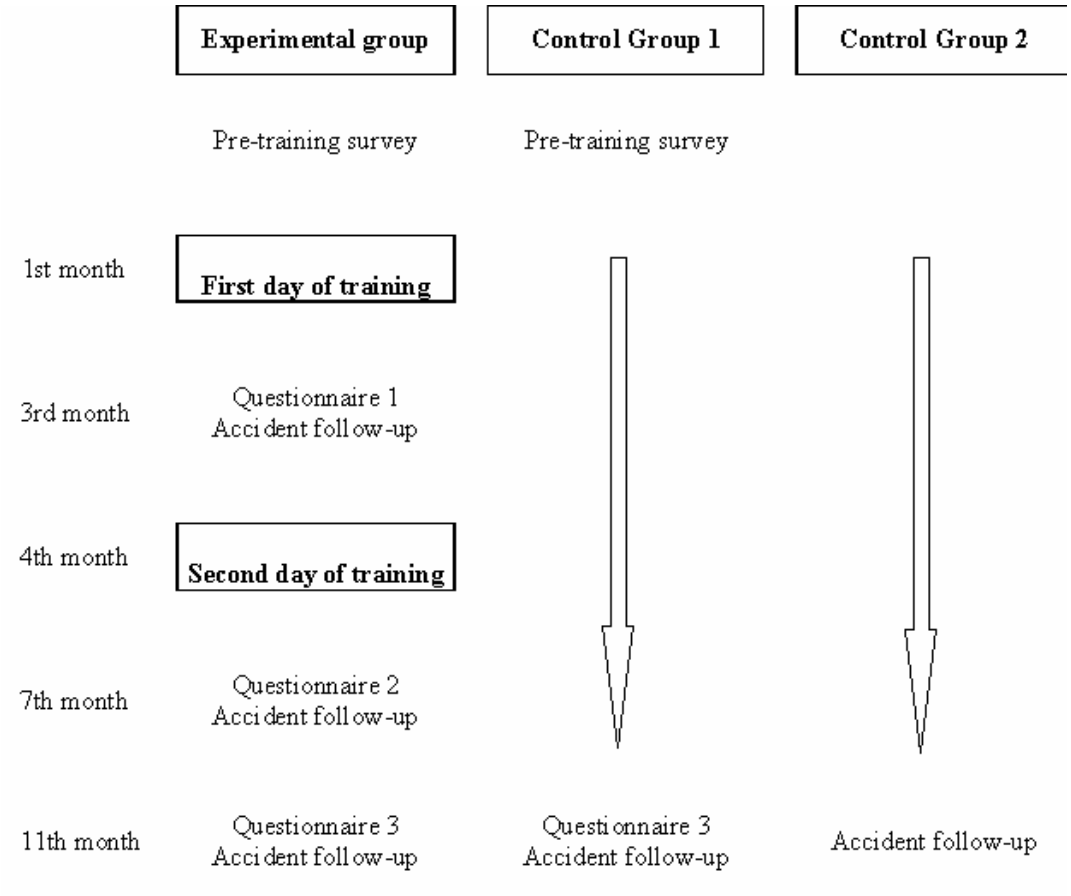
The drivers are from 7 French départements (administrative zones) across the Central West Atlantic region: Charente (16), Charente-Maritime (17), Loire-Atlantique (44), Maine et Loire (49), Deux-Sèvres (79), Vendée (85) and Vienne (86).



2.2 Procedure

3 groups of young drivers were monitored over a period of 11 months to assess if participation in a two-day training programme (experimental group) or if a pre-disposed orientation to road safety (experimental group and control group) had an influence on behavioural development and also on a reduction in terms of accident rates.

Table 2.2.1: The table below outlines the follow-up procedure of the various groups, as selected by ECF:



N.B. With regard to the experimental group, only the pre-training survey and questionnaire 3 (at +11 months) were analysed.

2.3 Allocation

586 young people agreed to participate in the project. They were then randomly assigned to two groups:

- **Experimental group** : this group followed the training days and responded to the questionnaires.
- **Control group** : this group responded periodically to the questionnaires

Ultimately, of the 396 positive responses to the invitation, 124 participants took part in the training days. This represents a participation rate of 31 %.

Moreover, of the 190 questionnaires sent to the control group, 87 persons replied. This represents a response rate of 46 %.

Without their knowledge, a 3rd group, called control group 2 and made up of 124 young drivers, was monitored in terms of accidents by the MACIF.

This table describes all 3 groups according their numbers, sex, age and initial (pre-licence) driver training.

Table 2.3.1 – Description of the groups:

	Number of persons	**Sex**	**Age**	**Initial training**
Experimental group	124	66 % male 34 % female	18 to 19 : 73 % 19 to 23 : 27 %	51 % Tradi* 49 % AAC**
Control group 1	87	41 % male 59 % female	18 to 19 : 52 % 19 to 23 : 48 %	51 % Tradi* 49 % AAC**
Control group 2	124	62 % male 38 % female	18 to 19 : 77 % 19 to 23 : 23 %	49 % Tradi* 51 % AAC**

*Tradi : traditional driver training (professional) at 18 years old
**AAC : Apprentissage Anticipé de la Conduite (accompanied driving), a training starting at 16 years old involving a period of accompanied driving with parents.

3. Training Programme

ECF's prior experience with post-licence training convinced us of the importance of spreading the training over 2 separate days with a period of independent driving in between. This format also allowed for deeper discussions, exchanges of experience and debate amongst participants. The training course fits into an educational continuum designed for road users. It allows for a training period spread over a number of months which encourages reflection amongst the participants.

In overall terms, the training programme includes information, experiences and discussions on different risks, whether objective or subjective. It alternates between class-based sessions, traffic and track. The two training days have a 4 month interval in between.

The sessions take account of the hierarchical levels of driver behaviour : vehicle manoeuvring, mastery of traffic situations, journey-based contexts and motives, and attitudes and goals for life.

Particular importance at this post-licence stage was attached to the two last levels (3 and 4) via the relation that the young driver has with his car and his approach to risks. Our general aim was to allow the young driver to generate positive attitudes with regard to road safety and in such a way to improve the lower levels of the hierarchy model too.

The training encouraged interaction between participants and experimenting with different driving situations. The two training days allow for hypotheses, as well as an opportunity to verify them and then to raise awareness, acceptance and taking into account other road users. They are also founded on knowledge of the different limits of man, the vehicle and the environment.

In this sense, the training reinforces, through the notion of sharing the road, the notion of citizenship and proper attitudes and behaviour of the driver.

3.1 First training day

From a pedagogical perspective, the participants have to be able to check the validity of the information given by the trainers. The group discussion sessions take place before and after the external sessions, in order to generate a dynamic effect within the group.

3.1.1 Classroom session

Road safety policy
The objective of this session is to highlight the relationship between man, the vehicle and his environment. Through this, the participants should become more aware of their habits, personal characteristics and attitudes regarding their driving style and road safety.
The general presentation on road safety allows participants to put their own driving, and driving of others, into perspective. This debate also sheds light on the main reasons for accidents.

3.1.2 Traffic-based exercises

Driving audit
The driving audit allows participants to evaluate themselves and to consider different ways of driving. It is carried out in groups of 6 participants with a training in a people carrier car. Each participant drives for approx. 20 minutes.
The main objective of this session is to improve participants' self-evaluation skills (decision-making, hazard perception…). The driving audit encourages awareness of one's strengths and weaknesses in difficult situations and demonstrates that there are differences in approach and ability between one road user and another.

3.1.3 Classroom session

Subjective illusions
This session is designed to raise awareness of the differences between vision and perception. The workshop stresses the fact that interpretations of a situation are often subjective and can vary from one person to another.

3.1.4 Classroom session

Drugs and alcohol
The main objectives of this module are to counteract incorrect understanding of alcohol and drugs in relation to road safety. It is important to take into account all the remarks made by the participants during the simulation (SIMALC alcohol simulator) and to challenge them based on their own experience and knowledge. The module focuses on the effects of alcohol and drug consumption.

3.2 Second training day

3.2.1 Traffic-based exercises

Driving audit
This audit focuses on a driving assessment with a view to adopting a safe driving style. Any evolution in driving style since the first day's audit can also be evaluated.

3.2.2 External exercise

Discussion around cars of participants
Participants split into pairs and, with the help of evaluation forms, information is noted regarding the car of the other participant. The idea is to increase awareness of the importance of a fully maintained car for safe driving. Some observations which may seem superfluous at first can actually reveal certain contradictions in the behaviour and values of the owner. For example, a car which is not properly maintained can be a financial problem for the owner. But what if the car is over equipped with gadgets which cost more than the basic maintenance of the car?

3.2.3 Classroom session

Group reflection
This part of the day is devoted to the attitudes of the participants with regard to the training. It is a moment of reflection, where the trainer acts as a moderator and does not provide answers. The participants need to raise questions and generate discussion.

3.2.4 Track training session

The braking module and the loss-of-control simulator on a moving car are designed to raise awareness of the limits of the individual drivers and the car. These are by no means exercises for improving manoeuvring techniques. They are, above all, a means to experience and experiment (I feel my limits) and to be confronted with reality (I see that I cannot do it), I understand why, I accept it and I accept the 'other' (a pedestrian, a cyclist, etc).

These experimentation modules are designed to reinforce the classroom work and to lead to a constructive exchange of opinions and experiences.

The main objective of these practical exercises is to understand speed and dynamic factors related to cars. The notion of speed needs to be learned first of all using technical and physical explanations. The linked subject of dynamics allows participants to understand how vehicles function on the road.

The braking exercise
The objectives are to raise awareness of the difficulties linked to an emergency braking manoeuvre and the limits of ABS. The method used in this exercise is a demonstration with commentary by the trainer.

Contents
The exercise deals with the whole question of ABS. The trainers ask the participants to think rationally about what ABS is really designed for. The trainer demonstrates a braking manœuvre at 50 kmh without ABS and at 70 kmh with ABS, and he is accompanied by a participant who checks the speed of the vehicle.

Aim of the exercise : Prior to the exercise, the participants are asked to estimate the braking distance from the braking point. It is important to let the participants decide on their own because they generally underestimate the braking distance.

To the general surprise of the participants, the braking distance doubles when raising the speed from 50 kmh without ABS to 70 kmh with ABS (illustrated using kinetic formulae).

The question is asked again : what is ABS good for ?
Generally, the response of the participants is : " It's rubbish !" - " It's useless !".

A further demonstration is made showing the braking distance at 50 kmh with ABS. The braking distance is seen to be either the same as without ABS or even longer.

After discussion amongst the trainees and trainers, the trainer carries out 2 braking and avoidance demonstrations at 50kmh :

- The first one is with ABS (pointing out that the braking distance increases considerably when carrying out an avoidance manœuvre)
- A second without ABS, with wheels blocked

The participants are then in a position to understand that ABS is an excellent tool for preventing the wheels blocking when braking and allowing for steering at the same time, but that ABS does not make braking more effective, in contrast to the expectations of the participants and to the claims of car salesmen.

The exercise reveals the incorrect assumptions about ABS and the implications of kinetic energy.

At each demonstration, the trainer specifies that the braking distance includes the distance covered during the reaction time.

The loss-of-control simulator
The objectives for this session are :
To accept the imperfections of the man, vehicle and environment triangle.
To raise awareness of the consequences of loss of control when driving.

Two methods are used to reach these objectives : the discovery method and practice with commentary.

Contents
The trainer lays the groundwork :
" You are going to experience the loss of control of a vehicle. "
An initial demonstration is given so that the participants can visualise the circuit.

The participants take it in turns to drive the car :
The trainer can be outside the vehicle in order to give a commentary on what is happening.
The vehicle is driven in first or second gear. After each participant has failed, the trainer asks the participants to think about what has happened.

After the trainees have made their suggestions on the different ways of avoiding loss-of-control, a demonstration is given on the basis of these suggestions.

The exercise also stresses the difficulty of recognising loss of control and thus the impossibility of regaining control of the vehicle (reaction time, width of the road…).

4. The Trainers

The entire training is dependent on the quality of the trainers and the educational quality. The 5 trainers are road safety professionals who are trained in awareness programmes for young drivers. They worked in pairs throughout this training. A training guide was formulated in order to provide additional structure and coherence to the programme. The objectives of the guide are based on the P.N.F. (*Programme National de la Formation à la conduite* = National Programme for Driver Training).

Whilst the guide was designed to ensure coherent implementation of the training, from one trainer to another, the methods used were determined by each pair of trainers in order to adapt them to the group. The trainer's guide lists the objectives of the training and the links between each sequence and module.

5. Feedback on the training

5.1 Feedback from organisers and trainers

The organisation of this two-day training programme generated considerable exchange and debate amongst the participants and required a major investment on their part.

The training programme over the two days seemed coherent because the questions raised after the first day were answered in the second day. The extra length (2 days) also allowed trainers to listen and encourage reflection more on each of the subjects addressed. The trainers themselves appreciated the quality of the different debates and workshops that took place.

Generally, the trainers' feedback is very positive. They are not frustrated because they can see the trainees a second time, share new experiences and observe some form of development with regard to road safety.

One idea may be to replace the second driving audit (second day) with a debate on how the participant's driving styles have changed in the interim period between the two training days.

5.1.1 Feedback on the classroom sessions

In our opinion, none of the subjects addressed in the classroom modules should be removed. However, if such an programme was to continue, an extra regulatory item would be added concerning the recently introduced provisional licence and how to avoid traffic offences.

The level of discussion generated between the participants and the trainers was high. The trainees reacted well. The discussions benefited from an informal and open setting without any form of moralising or prejudice.

The main visual aids used were overhead slides, a whiteboard and a flipboard.

The trainers all worked in the same location and thus aided feedback on the classroom sessions.

5.1.2 Feedback on the track exercises

Again, this session was rated positively. The braking exercise and the loss of control simulator allowed the participants to understand their own personal limits and the limits of the vehicle.

5.1.3 Feedback on the driving audit

The audits, which take place in groups of in a people carrier, allows each participant to reflect on his driving style and those of other drivers. We believe that this training should take place in a familiar location for the participants. This allows the participant to feel natural and for the evaluation to be more accurate. The length of 20 minutes driving per person allows enough time to carry out the analysis and to reflect on one's driving habits.

The trainers regularly compared their assessment forms to ensure a coherent approach to the evaluations. The trainers know how to keep a low profile during the audit so that the participant can drive as spontaneously as possible.

5.2 Feedback from participants

This data was collected from post-training surveys 1 and 2. Post-training survey 1 took place 3 months after the first training day (i.e. before the second training). Post-training survey 2 took place 3 months after the second training day.

Table 5.2.1 - The ECF/MACIF training is :
(from 1 = « not at all » to 5 = « Absolutely »)

At 3 mths: N=80	Useful	Interesting	Motivating	Valuable	To be repeated
Mean	3,9	4,3	3,7	4,1	4,0
Sd	0,8	0,6	0,9	0,9	1,0
At 7 mths: N=59					
Mean	4,4	4,5	4,2	4,5	4,4
Sd	0,6	0,5	0,8	0,7	0,8
P(z)	0,0001	0,05	0,01	0,01	0,02

The differences in the means between the first and the second measurement can easily be explained by the fact that the participants attending the 2nd training are clearly motivated to participate.

Table 5.2.2 - What do you recall above all from this training day ?
(from 1 = « not at all » to 5 = « Absolutely »)

At 3 mths: N=80	The good spirit	Technical aspects	Practical aspects	The theory
Mean	4,2	3,4	3,6	3,7
Sd	0,7	0,9	1,0	0,9
At 7 mths: N=59				
Mean	4,3	4,1	4,3	3,3
Sd	0,6	0,8	0,6	0,9
P(z)	ns	0,0001	0,0001	0,02

The differences in the means observed for the two days can be explained by the different content of each training day.

Table 5.2.3 - The trainers were : (from 1 = « not at all » to 5 = « Absolutely »)

At 3 mths: N=80	Advisors	Moralisers	Good listeners	Open	Dynamic	Convincing	Approachable
Mean	4,1	2,4	4,6	4,7	4,7	4,4	4,5
Sd	0,7	1,3	0,5	0,4	0,4	0,6	0,6
At 7 mths: N=59							
Mean	4,2	2,3	4,5	4,6	4,6	4,3	4,6
Sd	0,7	1,1	0,5	0,5	0,5	0,6	0,5
P(z)	ns	ns	ns	ns	ns	ns	ns

Table 5.2.4 - Since the training : (from 1 = « not at all » to 5 = « Absolutely »)

At 3 mths: N=80	I have acquired new technical knowledge	I have a different perspective of others	I adapt my speed to the situation in hand
Mean	3,2	3,6	3,8
Sd	1,1	1,1	0,9
At 7 mths: N=59			
Mean	4,1	3,6	4,1
Sd	0,9	1,0	0,8
P(z)	0,0001	ns	ns

Table 5.2.5 - Which were the most interesting themes dealt with in the training : (from 1 = « not at all » to 5 = « Absolutely »)

At 3 mths: N=80	Road Safety Objectives	The study on reaction times	Alcohol	Driving in traffic	Drugs	Tiredness	Braking	The loss of control simulator	The vehicle check
Mean	3,5	3,6	4,2	3,8	3,9	4,0	-	-	-
Sd	0,9	0,9	0,8	0,9	1,1	0,9	-	-	-
At 7 mths: N=59									
Mean	-	3,7	-	3,6	-	-	4,3	4,7	3,3
Sd	-	0,7	-	0,8	-	-	0,7	0,6	0,8
P(z)	-	ns	-	ns	-	-	-	-	-

No significant differences were found in the ratings of subjects dealt with on both training days.

Table 5.2.6 - Which subjects did you find the most useful : (from 1 = « not at all » to 5 = « Absolutely »)

At 3 mths: N=80	Road Safety Objectives	The study on reaction times	Alcohol	Driving in traffic	Drugs	Tiredness	Braking	The loss of control simulator	The vehicle check
Mean	3,5	4,1	4,4	4,1	4,2	4,4	-	-	-
Sd	1,1	0,9	0,8	0,8	0,9	0,7	-	-	-
At 7 mths: N=59									
Mean	-	4,3	-	3,7	-	-	4,7	4,5	3,5
Sd	-	0,8	-	1,0	-	-	0,5	0,8	1,0
P(z)	-	ns	-	0,02	-	-	-	-	-

Table 5.2.7 - You found the length of each module : (from 1 = « Appropriate » to 5 = « Inappropriate »)

At 3 mths: N=80	Statistics on road safety	Driving in traffic	Visual perceptions	Drugs	Alcohol	Braking	Loss of control simulator	The study on reaction times
Mean	2,3	2,3	2,3	2,4	2,3	-	-	-
Sd	1,2	1,0	1,2	1,2	1,3	-	-	-
At 7 mths: N=59								
Mean	-	2,1	-	-	-	1,9	1,7	2,1
Sd	-	1,1	-	-	-	1,1	1,1	1,1
P(z)	-	ns	-	-	-	-	-	-

6. The Evaluations

The main evaluation aim was to demonstrate if the training affected young drivers' attitudes towards road safety and consequently if this had an effect on accident frequency.

Because the training was limited in length, it seems difficult to ascertain the level of behavioural change and any consequences on accidents. As a result, ECF and MACIF have decided to extend their monitoring beyond the timeframe of NovEV, in order to check any change in accident rates.

The participants were classified according to sex, age, initial training, socio-professional status and place of residence. Their profiles also include information on the car driven (type, power…) and the use of the car (reasons for driving, frequency of driving…).

In order to measure behavioural developments as the months went by, each questionnaire asked the participants to evaluate their vision on road safety, driving habits and driving skills, and their conception of risk.

Each questionnaire contained eight questions :
1. With respect to your driving habits and skills, what factors could present a risk to you ?
2. Every driver is different and has strengths and weaknesses concerning driving. What are yours ?
3. You see cars as a symbol of….
4. A good driver is someone who….
5. What situations or factors annoy you when driving ?
6. In the following table, a series of driving situations are presented (estimate how often these situations occur to you) (level 2)
7. What is the greatest risk to you in life ?
8. An accident is often due to….

Each of the above questions consists of 8-14 suggestions/statements which may be repeated differently in order to check the coherence of the responses.
The choice of a monitoring period of 11 months was made in order to accurately assess the effects of the training over time.

7. Factorial analysis

Factorial analysis was carried out on the total 101 candidates in the two groups. This analysis enabled us to retain 4 measurement scales.

Scale 1 : Evaluation of risks linked to driving habits (8 items)
Question 1 : «When considering your habits and skills, what factors could constitute a risk for you ? »
8 statements :
1 Driving too fast
2 Falling asleep at the wheel
3 Desire to race in traffic
4 Lack of respect for safety margins
5 Carelessness
6 Excitement, irritation
7 Over-confidence
8 Alcohol consumption

Responses were scaled between 1 and 5 (from « Not at all » to « Absolutely »).

Scale 2 : Evaluation of risks linked to the frequency of specific driving situations (10 items)
Question 6 : « Estimate how often the following occurs to you : »

10 statements :
1 I have driven while drunk
2 I have driven when nervous
3 I have driven when irritated
4 I have driven after having taken drugs
5 I have driven over the speed limit
6 Another driver challenged me in traffic
7 I have driven with a hangover
8 I have driven in order to calm down
9 I have followed another car too closely
10 I have driven too quickly for the circumstances

Responses were scaled between 1 and 5 (from« Never » to « Very often »).

Scale 3 : Evaluation of skills for defensive driving (5 items)
Question 2 : « What are your strengths and weaknesses with regard to driving ? »
5 statements :
1 I take into account pedestrians and cyclists
2 I respect traffic rules while driving
3 I drive carefully
4 I give way where necessary
5 I drive at an appropriate speed

Responses were scaled between 1 and 5 (from « Not at all » to « Absolutely »).

Scale 4 : Representation of risks (13 items)
Question 8 : "An accident is often due to"

13 statements :
1 A mechanical problem
2 Alcohol
3 Tiredness
4 Men
5 Women
6 Young people
7 Old people
8 Drugs
9 Speed
10 Bad weather
11 Lack of attention
12 Road conditions
13 Non-respect of rules

Responses were scaled between 1 and 5 (from« Never » to « Very often »).

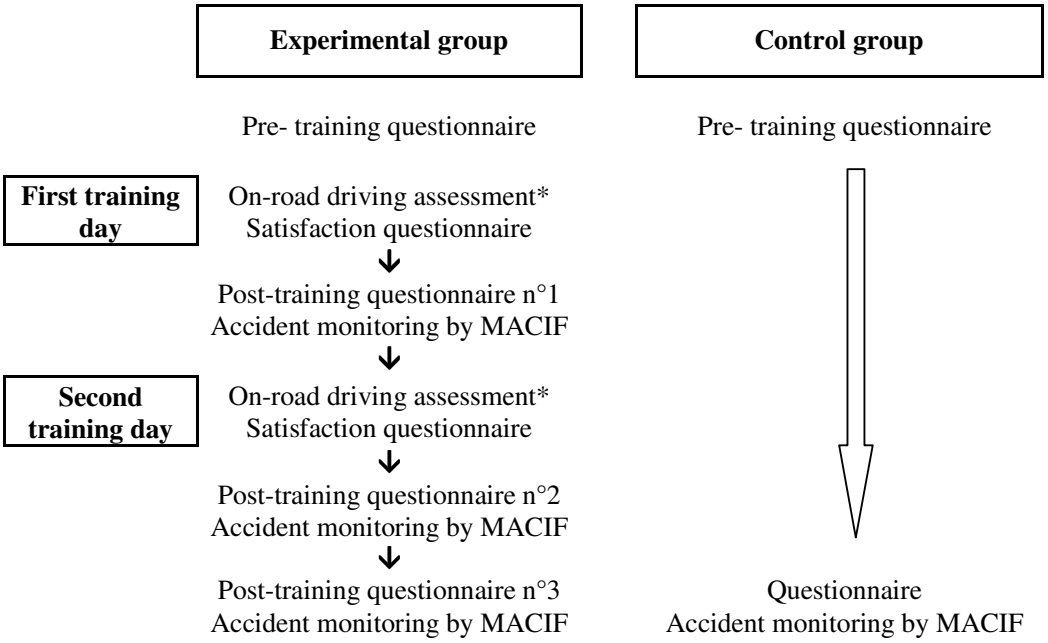
Table 7.1 – Cronbach's Alpha coefficients :

	Alpha de Cronbach
SCALE 1	0,92
SCALE 2	0,80
SCALE 3	0,82
SCALE 4	0,76

8. Data Collection Methods

The table below describes the evaluation methods and includes the time periods when the data was collected.

Figure 8.1 – Data collection process:



*The on-road driving assessment forms were used as pedagogical support for the feedback to participants, and between trainers and the organisers. They were not used in the statistical analysis.

The data was collected using different methods :

- The driving assessment forms and the satisfaction questionnaires were given directly to the trainers by the participants.
- The pre- and post- training questionnaires were sent to the homes of the participants. The participants then filled out the questionnaires and returned them to ECF in the enclosed stamped, addressed envelope.

9. Statistical methods

9.1 Data processing

The data from the different questionnaires were entered into an Access database. Each questionnaire was coded based on name, first name, birthday, etc. The questionnaires were then recoded to ensure anonymity and to check for doubles.

Table 9.1.1 - Total number of participants processed :

	Initial phase	**Final phase**	**Loss**
Experimental group	« Pre- training » Questionnaires 127	« Post- training » Questionnaires 67	47 %
Control group	Questionnaire « 1 » 93	Questionnaire « 2 » 51	45 %

The candidates selected for the study were present during the 2 phases of the test in order to measure evolution before and after the training:

Table 9.1.2 - Participants present during the 2 phases of the test :

Groups	**Number of candidates selected**
Experimental	61
Control	50

9.2 Cleaning the data

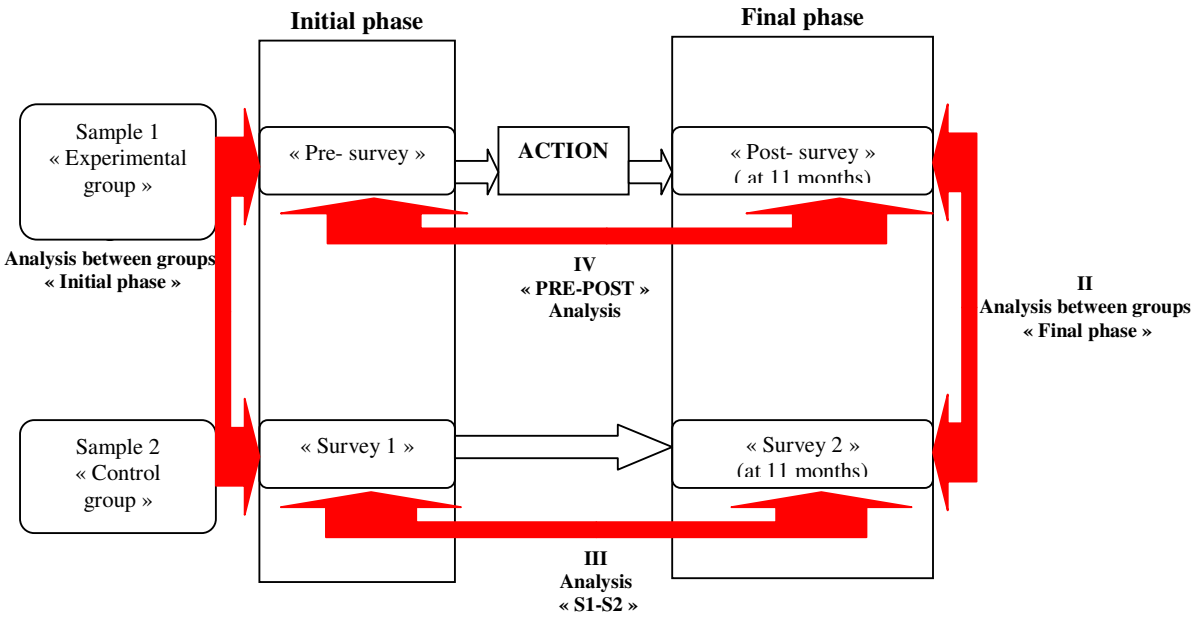
Firstly, the questionnaires which were less than 90% complete were eliminated. Unanswered questions were entered as 0 when the value was expected to be on the scale of 1-5 and when 0 was not a possible response.
Other questionnaires with illogical responses or answers which were too subjective were also eliminated. To do this, tests were carried out on pairs of questions. Average differences above 2,75 in absolute terms (with the response scale as always at 1-5) justified elimination of the corresponding questionnaire.

Table 9.2.1 - Participants after the cleaning (remaining for the analysis) :

Groups	**Number of candidates remaining**
Experimental	56
Control	45

9.3 Planning the analyses

Figure 9.3.1 – The 4 steps of analysis:



9.4 Hypotheses

I The two samples are similar at the beginning.
=> Analysis between groups « Initial phase »

II The two groups are different at the end.
=> Analysis between groups « Final phase »

III The control group has not evolved greatly between the two survey points.
=> Survey 1 - Survey 2 Analysis »

IV The experimental group shows significant changes when comparing before and after the training.
=>Analysis « Pre- survey - Post-survey »

For each part, the global results were compared, then for each specific category of question :

Scale 1 : Evaluation of risks linked to driving habits (8 statements)
Scale 2 : Evaluation of risks linked to the frequency of certain traffic situations (10 statements)
Scale 3 : Evaluation of defensive driving skills (5 statements)
Scale 4 : Representation of risks (13 statements)

At each level of analysis, the global results were compared, followed by results according to the criteria below :

« Sex » criteria : differences between males and females.
« Initial training » criteria : differences between traditional driver training and accompanied driving (AAC).
“Age” criteria: differences between 18, 19 and above 20 years old candidates.

10. Analysis and results

10. 1 Description of the groups

The analyses were based on the persons following the entire initiative. The numbers of the two groups are as follows :

Table 10.1.1 – Number of subjects :

	Number
CTRL	45
EXP	56
Total	101

NB. CTRL : Control group
EXP : Experimental group

Gender :

The sex distribution in the two groups differed significantly (Khi² - $p<0.05$). The experimental group was largely male and the control group contained mostly females.

Table 10.1.2 – Repartition by sex :

	Female	Male	Total
CTRL	26 58%	19 42%	45 100%
EXP	19 34%	37 66%	56 100%
Total	45 45%	56 55%	101 100%

Age :

The spread of age is unequal from one group to the other. The experimental group is younger than the control group. The difference between the two groups is significant (khi² - $p<0.05$) for the age groups 18 and 19 years old. There are no significant differences for the other age categories. The number of 20 year olds and above was very small so this category was grouped together.

Table 10.1.3 – Repartition by age :

	18	19	20	21	22	23	Total
CTRL	22 49%	17 38%	3 7%	0 0%	1 2%	2 4%	45 100%
EXP	43 77%	8 14%	1 2%	1 2%	2 4%	1 2%	56 100%
Total	65 64%	25 25%	4 4%	1 1%	3 3%	3 3%	101 100%

Initial training :

The distribution of this factor is homogeneous across the two groups.

Table 10.1.4 – Repartition by initial training:

	Tradit.	AAC	Total
CTRL	26 58%	19 42%	45 100%
EXP	30 54%	26 46%	56 100%
Total	56 55%	45 45%	101 100%

10.2 Inter group analysis – Initial phase

Fisher Tests : analysis of differences between groups with a confidence interval of 95,00 %

Scale 1 : Evaluation of risks linked to driving habits (8 items)
Question 1. according to your habits and skills, what factors could represent a risk to you :
Driving too quickly
Falling asleep at the wheel
A desire to race in traffic
Lack of respect for proper safety margins
Carelessness
Overexcitement, irritation
Excess confidence
Alcohol consumption

Difference between groups

There were no significant differences when both groups are considered as a whole. However, there were significant differences with regard to the women in the two groups, where the risk awareness level differed considerably from one group to the other ($p<0.002$). The risk awareness level is thus considerably lower amongst the women in the experimental group (figure 10.2.1).
The same observation ($p<0.02$) was made with regard to participants aged 20 and above (figure 10.2.2). This difference was primarily due to the women aged 20 and above ($p<0.003$).
Significant differences could also be observed concerning women aged 19 ($p<0.05$).
As far as the type of initial training is concerned, there were differences in the participants following traditional training: these levels were lower ($p<0.001$) in the experimental group (figure 10.2.3).

Figure 10.2.1 – Interactions between "group" and "sex" factors:

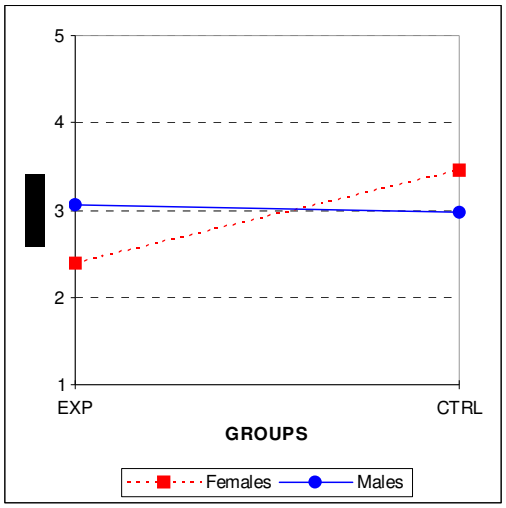


Figure 10.2.2 – Interactions between "group" and "age" factors:

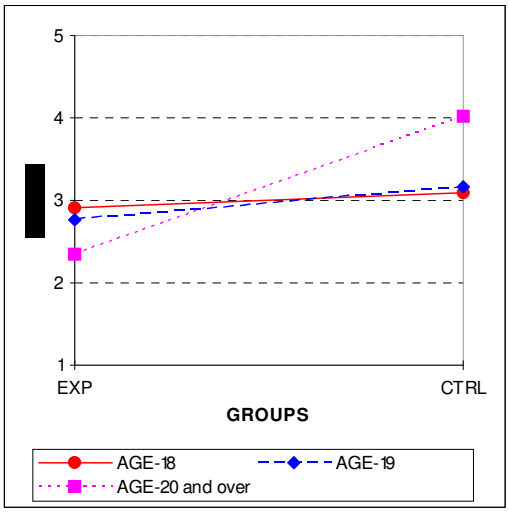
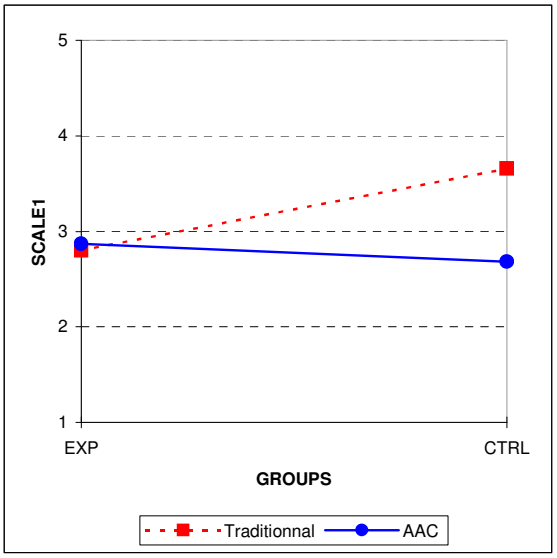


Figure 10.2.3 – Interactions between "group" and "initial training"



Scale 2 : Evaluation of risks linked to the frequency of specific driving situations (10 items):
Question 6. Estimate the frequency of these situations occurring :
I have driven while drunk
I have driven while worried
I have driven when annoyed
I have driven after having consumed drugs
I have driven over the speed limit
Another driver challenged me in traffic
I have driven with a hangover
I have driven to calm down
I have followed the preceding car too closely
I have driven too quickly according to the circumstances

Differences between groups :

Fisher Test : analysis of differences between groups with a confidence interval of 95,00 % :

The participants in the experimental group achieved lower results than those in the control group ($p<0.0001$ – figure 10.2.4). These differences were particularly amongst males ($p<0.001$) but were also observed to a lesser degree amongst females ($p<0.05$) (figure 10.2.5).
As far as age is concerned (figure 10.2.6), differences in level were particularly noticeable in the 18 yrs old category ($p<0.0001$) and the 19 yrs old category ($p<0.003$). These differences were mostly due to men in the 19 yrs old category.
As far as initial training was concerned (figure 10.2.7), the differences concerned the two types of training ($p<0.001$ for AAC – $p<0.006$ for traditional training). These differences are primarily due to males.

Figure 10.2.4 – Difference between groups :

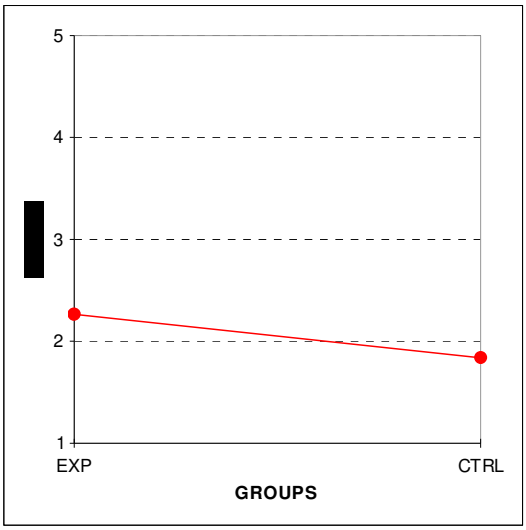


Figure 10.2.5 – Interaction between « group » and « sex »:

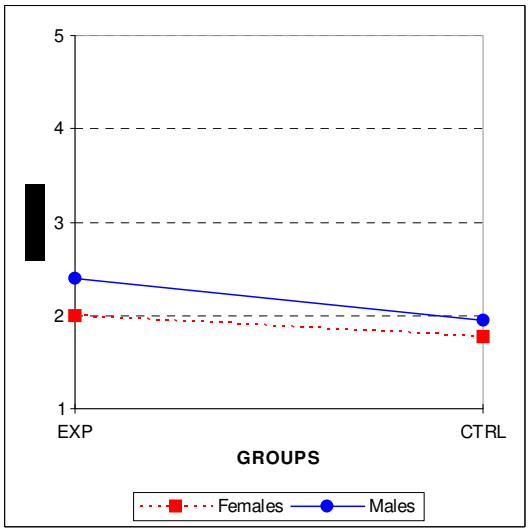


Figure 10.2.6 - Interaction between « group » and « age »:

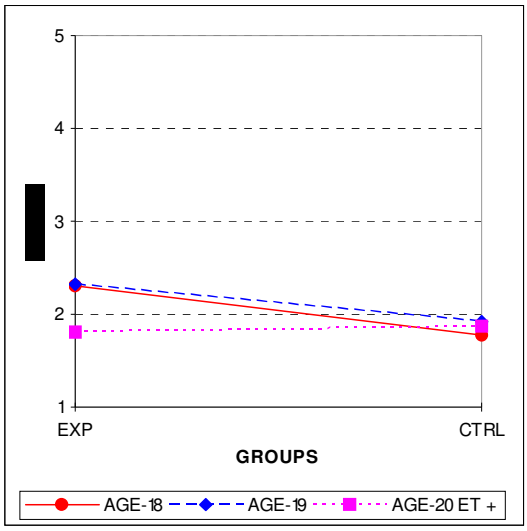
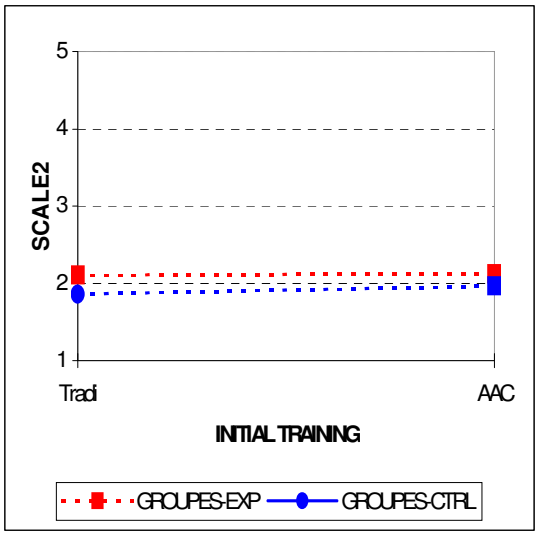


Figure 10.2.7 - Interaction between « group » and « formation »:



Scale 3 : Evaluation of defensive driving skills (5 items) :
Question 2. What are your strengths and weaknesses :
I take into account pedestrians and cyclists
I drive according to the traffic rules
I drive carefully
I give way when necessary
I drive at an appropriate speed

Differences between groups

Fisher test : Analysis of the differences between the groups with a confidence interval of 95,00 % :

On this scale, the participants in the experimental group achieved lower results than those in the control group ($p<0.0001$ – figure 10.2.8). These differences are as present amongst males ($p<0.02$) as amongst females ($p<0.005$ - Figure 10.2.9).
When considering the age of the candidates, the same trend was observed amongst the 18 and 19 yr old categories. In these categories, it was mostly the males who accounted for the difference (figure 10.2.10).
The differences were also visible regardless of training types (figure 10.2.11). They were due to the males who followed AAC and to the females who followed traditional training.

Figure 10.2.8 – Difference between groups :

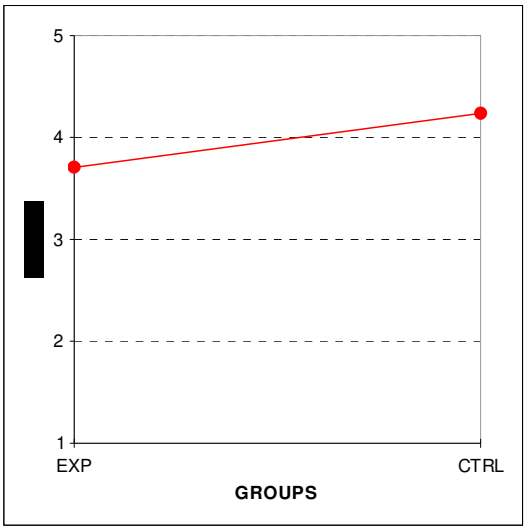


Figure 10.2.9 – Interaction between « group » and « sex »:

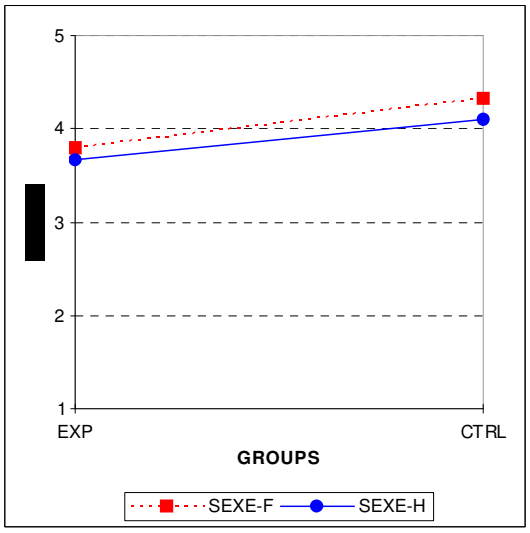


Figure 10.2.10 - Interaction between « group » and « age »:

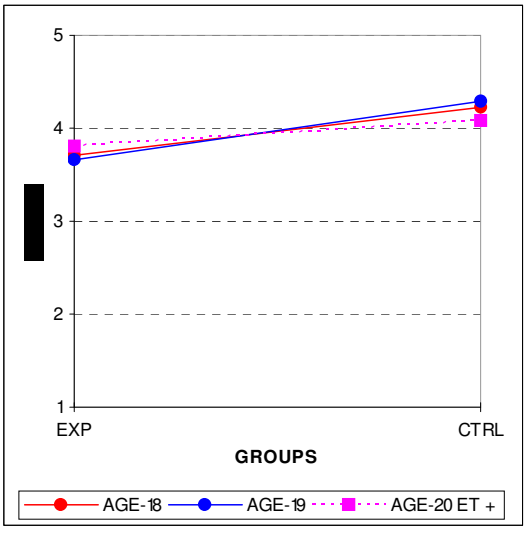
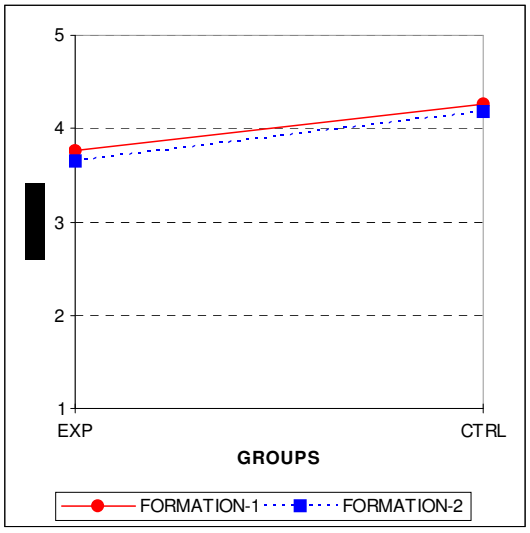


Figure 10.2.11 - Interaction between « group » and « formation »:



Scale 4 : Representation of risks (13 items) :
Question 8. An accident is often due to :
A mechanical fault
Alcohol
Fatigue
Men
Women
Youngsters
Old people
Drugs
Speed
Bad weather
Lack of attention

State of the road
Non-respect of rules

No significant differences.

Inter group analysis – initial phase

The original hypothesis was that the two groups (experimental and control group 1) are similar at the outset. However, in 2 of the 4 scales we can see that this is not the case. The experimental group shows a tendency towards more risky behaviour. This is confirmed when looking at scale 3 where the experimental group shows less aptitude for defensive driving.

10.3 Inter group analysis – Final phase

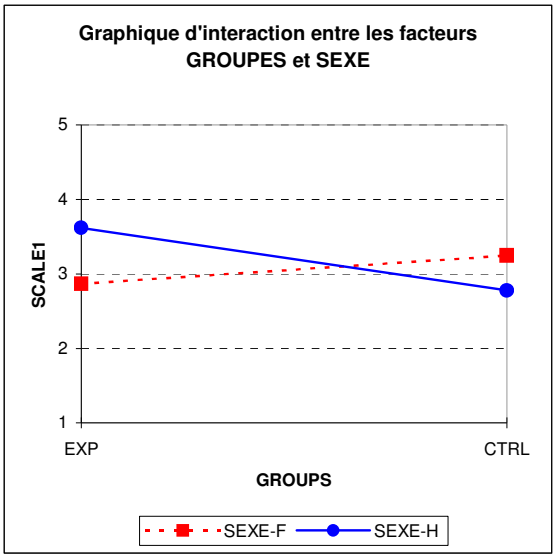
Scale 1: Evaluation of risks linked to driving habits (8 items)

<u>Differences between groups :</u>

Fisher test – Analysis of differences between groups with a confidence interval of 95,00 % :

When considering both groups as a whole, there are no significant differences. There are, however, significant differences amongst males ($p<0.005$ – figure 10.3.1). The males from the experimental group obtained better results in terms of risk awareness than those in the control group. This development is particularly noticeable amongst males in the 19 yrs old category ($p<0.01$).
In terms of initial training, the only observations with significant differences involved males who followed AAC training whose results were better in the experimental group ($p<0.05$).

Figure 10.3.1 – Interaction between « group » and « sex »:



Scale 2 : Evaluation of risks linked to the frequency of specific driving situations (10 items)

Differences between groups :

Fisher test – Analysis of differences between the groups with a confidence interval of 95,00 % :

The differences observed in the initial phase were still visible here. The results of the experimental group remained lower than those in the control group ($p<0.02$), albeit with a reduction in the gap (figure 10.3.2).
The differences here involved participants who followed traditional training ($p<0.04$ – figure 10.3.3), particularly the males who followed such training ($p<0.03$).

Figure 10.3.2 – Difference between groups :

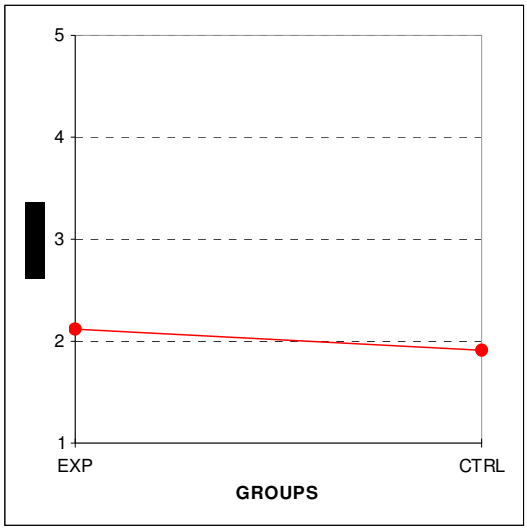
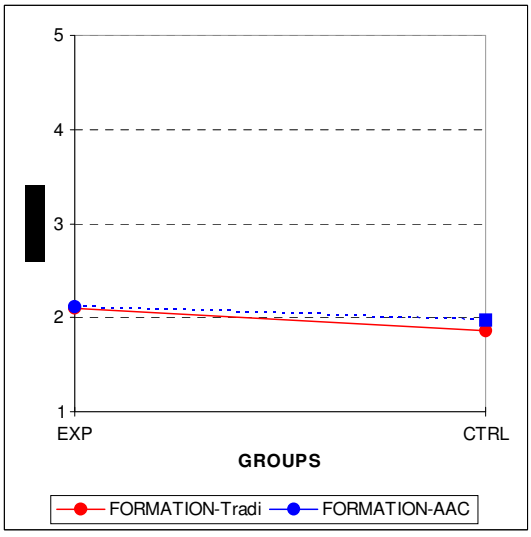


Figure 10.3.3 – Interaction between « group » and « formation »:



Scale 3 : Evaluation of defensive driving skills (5 items)

Differences between groups:

Fisher test – Analysis of the differences between groups with a confidence interval of 95,00 % :

The differences observed in the initial phase persist, although the gap between the two groups has reduced. The experimental group still scored lower in terms of a defensive driving style ($p<0.002$ – Figure 10.3.4). It was mostly the females in the experimental group who account for these results ($p<0.005$ – Figure 10.3.5).
In terms of initial training, the differences are significant for the participants who followed traditional training and those who followed AAC ($p<0.04$ – Figure 10.3.6). For the latter, it was women who account for the differences observed ($p<0.02$).

Figure 10.3.4 – Difference between groups :

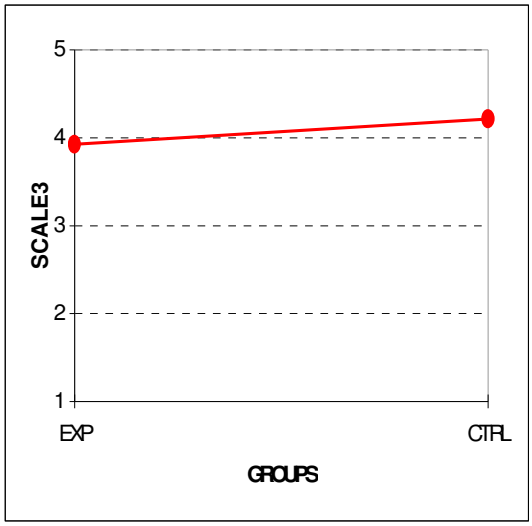


Figure 10.3.5 – Interaction between « group » and « sex » :

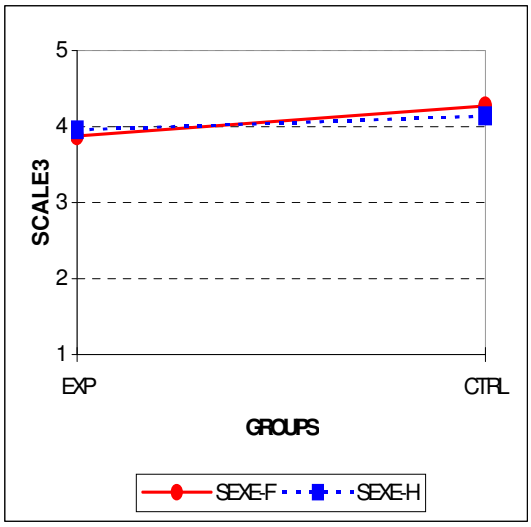
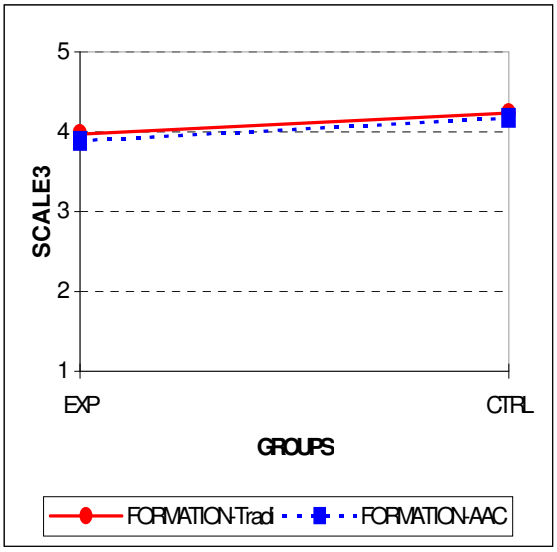


Figure 10.3.6 – Interaction between « group » and « formation » :



Scale 4: Representation of risks (13 items)

<u>Differences between groups :</u>

Fisher test- Analysis of differences between groups with a confidence interval of 95,00 % :

No significant differences.

Inter Group analysis – final phase

The differences observed in the initial phase are still present. However, it appears that the gaps have become smaller as far as scale 2 is concerned, as shown by the probability values. The same applies, although to a lesser degree, to the results of scale 3.

Conclusions on the inter group analysis

We can make no conclusions at this stage in any direction without knowing the development of each of the two groups.

10.4 Analysis – Control group

Analysis of control group

The analysis showed no significant differences in the scores of any of the scales of the control group.
The analysis of the different variables also revealed nothing, and as such the original hypothesis concerning the control group can be confirmed.

10.5 Analysis – Experimental group

Scale 1 : Evaluation of risks linked to driving habits (8 items)

There were positive changes observed concerning risk awareness linked to driving habits ($p<0.005$ – Figure 10.5.1). These changes were mostly due to the males ($p<0.02$). The differences seen between males and females (figure 10.5.2) before the training ($p<0.02$) are, naturally, more pronounced after the training ($p<0.008$) as a result of the evolution amongst males.
Results concerning initial training revealed positive changes amongst participants who followed traditional training ($p<0.04$ – Figure 10.5.3), especially amongst females ($p<0.05$). The males who followed AAC changed significantly for the better ($p<0.03$).

Figure 10.5.1 – Difference between phases:

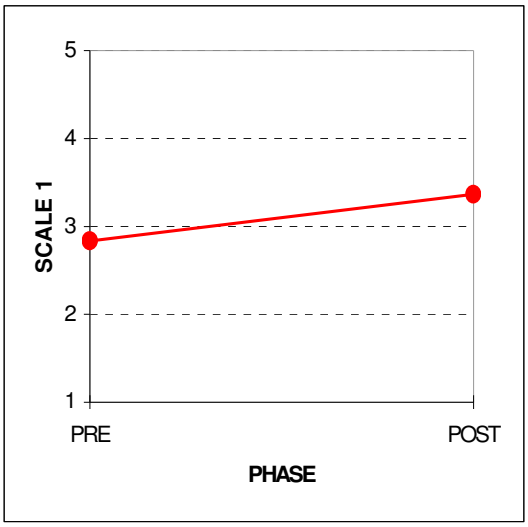


Figure 10.5.2 – Interaction between "phase" and "sex":

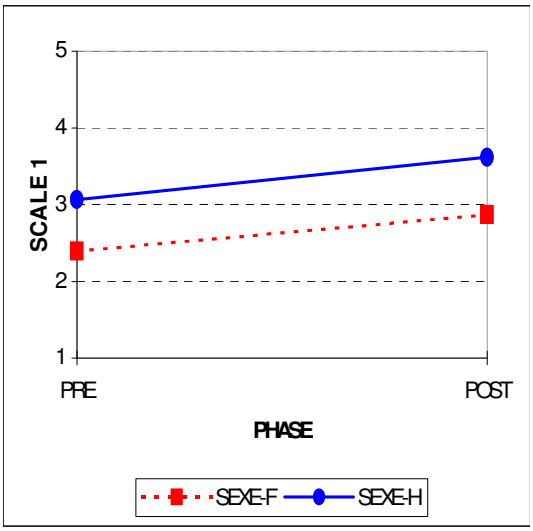
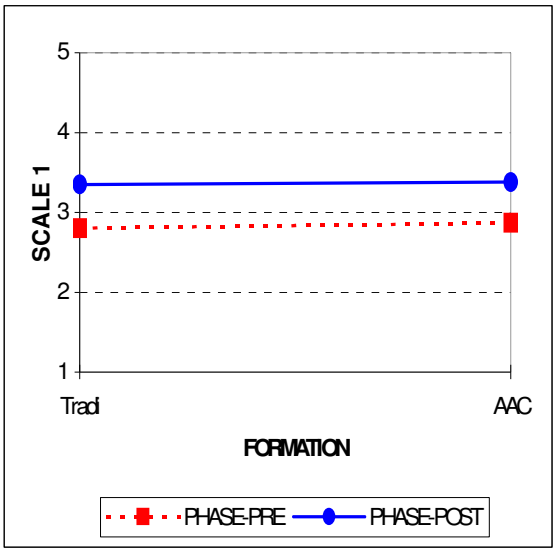


Figure 10.5.3 – Interaction between "phase" and "training":



Scale 2 : Evaluation of risks linked to the frequency of specific driving situations (10 items)

No significant differences were observed overall.
Only males who followed AAC initial training showed positive changes on this scale ($p<0.03$), in that they report a reduction in the frequency of risky driving situations after the training.

Scale 3 : Evaluation of skills for defensive driving (5 items)

There was a significant positive change overall concerning skills for defensive driving ($p<0.05$) (figure 10.5.4). This change is due to the males ($p<0.03$ – figure 10.5.5).

Figure 10.5.4 – Difference between phases :

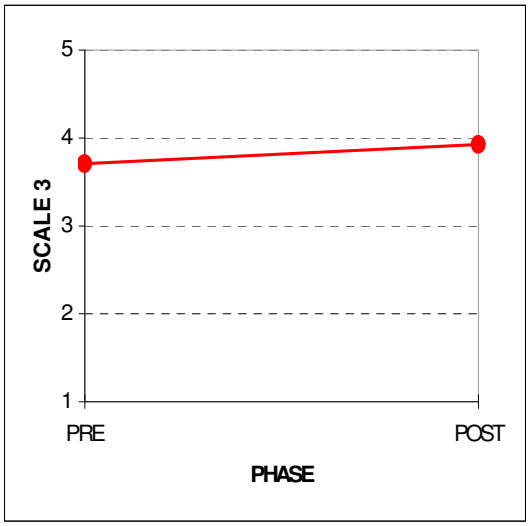
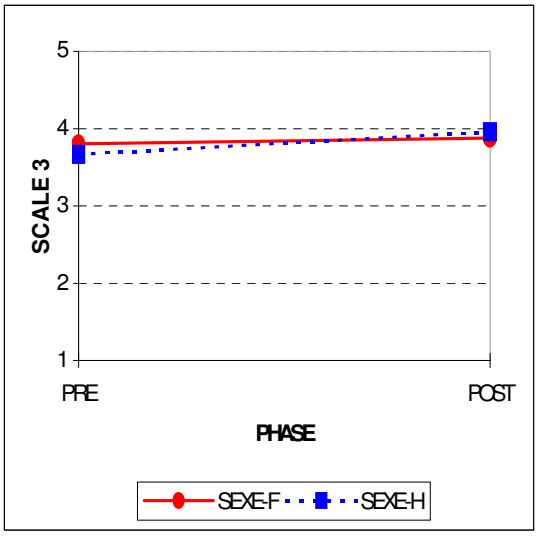


Figure 10.5.5 – Interaction between « phase » and « sex »



Scale 4: Representation of risks (13 items)

No significant differences.

Analysis of experimental group
The experimental group shows significant different results as far as scales 1 and 3 are concerned.

With regard to the evaluation of risks linked to driving habits (scale 1), the significant rise in means seems to demonstrate an increased awareness of risk factors.
The results which were below those of the control group in the initial phase have rejoined and passed the values of the control group in the final phase.
This is due as much to the increase in the results of the experimental group as to the decrease in those of the control group (the latter are not statistically significant).

The other significant result in this group concerned the evaluation of defensive driving skills (scale 3). Here again, this significant improvement in results seems to indicate a better understanding of traffic risk.

Finally, to a lesser degree, a positive effect was observed concerning evaluation of risks linked to the frequency of driving situations (scale 2) amongst males.

11. Accidents

The following data was provided by ECF's partner, the MACIF insurance company.

	Experimental Group	**Control Group**	**Control Group 2**
Frequency	2 out of 124	2 out of 87	17 out of 124
Percentage	1,6 %	2,3 %	13,7 %
Sex	2 males	1 male 1 female	8 males 9 females
Initial training	2 AAC	2 Tradi	10 Tradi 7 AAC

A comparison in terms of accidents of the experimental group and control group 2 is significant. These two groups are composed of 124 persons who are virtually homogeneous in terms of their profile: 66 % males in group 1 compared to 62 % in group 2. Over a monitoring period of 11 months, 2 participants in group 1 were responsible for accidents as opposed to 17 in group 3. However, the two groups are not comparable because the experimental group – who volunteered to take part in the training – are likely to be considerably more safety-oriented than control group 2 who did not participate in the training and who were monitored without their knowledge by MACIF.

12. Conclusions

The two groups (experimental and control groups) showed differences at the beginning of the study (hypothesis 1). These differences seemed to be structural ones (the most obvious one being the distribution of sexes in the two groups). Besides, the safety levels of the two groups at the beginning of the study are in favour of the control group.

The control group did not go through any significant changes between the initial and final phases. This hypothesis is thus confirmed.

The experimental group saw some positive changes, as has just been shown.

We can reasonably conclude that the development of the two groups shows an increase in risk awareness in the experimental group. This helps to delay the phenomenon of overconfidence which is so often observed amongst novice drivers. Otherwise, the control group, which was followed statistically but not involved in the training, remained stable in its results, despite its clear investment in road safety (by wanting to take part).

This confirms the pedagogical approach taken by ECF and MACIF over the last 10 years. In fact, we consider it vital to ensure an ongoing continuum of action in the field of road safety awareness throughout one's life. This approach remains positive whatever the exact nature of the action (training, surveying, campaigns..) in the sense that it calls into question the relationship that the driver has with regard to his vehicle and to risk. It would also seem logical that, in the absence of training – or with training which is wrongly based on skills and mastery of vehicle control-, we can only expect to reinforce the process of overconfidence amongst novice drivers which is naturally acquired.

- GERMANY

NovEV
Results of the German second phase evaluation

Frank Prücher

Author: Dr. Frank Prücher (BASt, Germany)

Evaluation of the German "Voluntary Post-Licence Training Courses for Probationary Drivers"; FSF-model.

1 Summary of project

In April 2003, a voluntary second-phase training programme for probationary (novice) drivers (FSF) was introduced in Germany by law as a pilot project. Between 2003 and 2010, the FSF project aims to evaluate how, if at all, it contributes to reducing novice drivers' accident risk. To date, 13 out of 16 federal states in Germany have joined the pilot project and have started the training which offers an incentive in the form of a one-year reduction of the probationary period for the novice drivers who participate. The FSF courses actually started in spring / early summer 2004.

Before then, some preparatory work had already been carried out by the DVR (Deutscher Verkehrssicherheitsrat = German Road Safety Council). The DVR developed the manuals and subsequently trained the trainers to coach the seminar leaders (driving instructors). When the training was introduced in practice, approximately 1,500 seminar leaders and 200 track instructors had already been trained to implement the FSF model. In the first five months of training, about 200 novice drivers took part in the FSF-courses. The training is composed of five sessions, including three group discussion sessions, one 'training and observation' drive on public roads, and a track-based training programme. Overall, the demand for the courses within the target group of young drivers has been rather low so far.

BASt collected the addresses of all people involved in designing and implementing the FSF programme so far. Six quasi-identical questionnaires were developed for the six groups and sent to the persons involved: programme authors, the trainers of the trainer, the seminar leaders and track instructors, and the participants (novice drivers). Data collection took place in June and July 2004. Due to the tight deadline, no pre-testing or follow-up-testing could be performed. The main focus of the evaluation was on the perception of the training itself among the different categories of people involved and on to what extent the programme is transferred to the participants in such a way the programme authors intended it.

The German evaluation project was a process evaluation with a single measurement. Different views on the programme with regard to the organisation, implementation and achievements were collected from the six different groups involved, such as the course designers and the participants of the programme. The effectiveness of the programme for the young drivers was also analysed by comparing the learning goals set down in the manual with reported self-assessment and an evaluation of the participants' success in implementing the goals of the training in practice.

The results of the study provide an indication of how accurately the FSF training concept was implemented in practice. The results show that most of the programme is being performed as the authors intended. The task reports and ratings of the importance of the programme modules correspond closely to the authors' concept and manual, the implementation as carried out by the trainer, and the participants' experience with FSF. Similar statements on the FSF modules were found in the six groups surveyed. The training was generally rated positively by the participants. The participants claimed to have developed and used several intellectual and behavioural strategies for safe driving, which is the main aim of FSF.

Attention should be turned to the fact, that participants reported an unwanted improvement of their abilities to master difficult traffic situations. Also the training of the track instructors should be revised because they perceived and implemented the track training course with other goals and intentions than was specified by the authors.

Further efforts to motivate novice drivers to participate in FSF should be made. Suggestions on how to do this are made in the conclusions.

2 The German evaluation project

2.1 Research questions, evaluation design, and objective of the project

The German evaluation of the FSF-model was carried out as a sub-project of the EU NovEV project. In 2002, when the NovEV project started, Germany already had a long tradition of post-licence training for drivers. A range of such measures had already been established and broader experience was available in addition to some research results. A brief overview of the existing measures and relevant research results will be given below.

In the proposal for the German sub-project, an evaluation of the effectiveness of one of the German second phase measures was proposed. It turned out that the EU and the NovEV project leader were mainly interested in the evaluation of a new second-phase model, the FSF-model, which is being introduced in Germany at the moment. This model is supported by important road safety organisations and interest groups (among others: DVR Deutscher Verkehrssicherheitsrat e.V. / German Road Safety Council, BFV Bundesvereinigung der Fahrlehrerverbände / Federation of Driving Instructor Associations, and ADAC Allgemeiner Deutscher Automobilclub / General Automobile Club of Germany).

After discussions with CIECA and in the NovEV project-team, an agreement was made to choose the FSF-model for evaluation in the German sub-project. 'FSF' stands for “Voluntary Post-Licence Training Courses for Probationary Drivers” ("Freiwillige Fortbildungsseminare für Fahrerlaubnisinhaber auf Probe").

The FSF-model contains all characteristic elements of second-phase programmes: classroom-training, an on-road feedback drive and track training. An important feature of the programme is that participation leads to a privilege in one's driver status: completion of the course leads to a one year reduction in the probationary period of 2 years for novice drivers in Germany. (As a result, legal regulations were necessary before the introduction of the model).

As the legislative process is unpredictable, and the introduction of the model was controversial, it was not possible to predict exactly the implementation date of the training. This constituted a major problem for decision-makers in the German sub-project, particularly with regard to questions concerning the research methods and evaluation design.

The decision to evaluate the FSF-model was taken when it was quite sure that it would be introduced early enough to make evaluation work possible within the timeframe of the NovEV project. As the practical introduction took place in the middle of 2004 – after the legal introduction at the end of 2003 and the beginning of 2004 – it was no longer possible to follow a classical before-after-design with a control group and a look at effects. As a consequence, the evaluation focus was shifted from questions of effectiveness to questions of pedagogical quality achieved in the practical implementation of the training. In this kind of process evaluation the focus is on the consistency of the actors’ empirical understanding of the model's contents – internally between the persons involved, and externally in relation to theoretical reference models like the GDE matrix (“good driver hierarchical model”, see Hatakka, Keskinen, Gregersen & Glad, 1999). Such a process evaluation was considered appropriate and important by the CIECA project management.

Empirically-supported responses to these questions (questionnaires were designed and sent to all relevant categories of persons involved in FSF) will hopefully support the quality development of the FSF model in many pedagogical and organisational aspects related to the work with young drivers and with the pedagogical staff working with the FSF model.

2.2 Participants/subjects/sample

Six different groups participated in the study. All members of these groups were involved in the development or implementation of the programme. BASt collected the addresses of all persons involved and questionnaires were sent out in June 2004. Altogether 970 questionnaires were mailed. The six groups are (number of questionnaires sent in brackets):

1. authors, consultants, and multipliers of the programme (N = 11)
2. trainer of the seminar leaders (N = 82)
3. trainer of the track instructors (N = 5)
4. seminar leaders (N = 487)
5. track instructors (N = 245)
6. participants (N = 140)

All questionnaires that were returned to BASt before August, 11th 2004 are considered in the sample. Altogether, 53.9 % of all questionnaires were filled out and returned. These 523 questionnaires form the final sample. Table 32 gives an overview of the basic characteristics of the whole sample.

Table 32: Rate of return and distribution of sex and age in the six groups

Group	% returned	Sample N	Sex				Age		
			Male		Female		Years		
			N	%	N	%	N	X	SD
authors / multipliers	72,7	8	-	-	-	-	-	-	-
trainer of seminar leaders	58,5	48	43	91,5	4	8,5	47	50,0	9,5
trainer of track instructors	100	5	4	80	1	20	5	46,8	5,5
seminar leaders	50,5	246	229	95,8	10	4,2	238	48,3	8,7
track instructors	59,6	146	140	95,9	6	4,1	143	45,9	8,8
participants	50	70	43	61,4	27	38,6	70	20,7	5,1

The next chapters provide a detailed analysis and some background information on the six groups based on the information collected from the questionnaires.

2.2.1 Sample characteristics participants

The average age of the 70 training participants was about 21 years with a range from 18 to 50 (only two of all participants were older than 27 years, the median age of the participants was MD = 19). About 61 % of all participants were male. Most of the participants came from the federal state of North Rhine-Westphalia (34.3 %), followed by Saxony (30.0 %), Hesse (14.3 %) and Baden-Württemberg (11.4 %); the remaining participants came from Bavaria, Thuringia and Saxony-Anhalt (less than 5 %).

The participants were mainly trainees (60.3 %), others were fully employed (19.1 %) or students (7.4 %). 5.9 % were currently unemployed and 4.4 % were doing their military or civil service. Table 33 shows the number of attempts taken to pass the driving test and the incidence of accidents before participating in FSF, according to gender in the participant group.

Table 33: Attempts at driving test and accidents before FSF participation

	Male	Female	All	
	N	N	N	%
Pass theoretical driving test				
1st attempt	36	22	58	82,9
2nd attempt	4	3	7	10,0

	Male	Female	All	
	N	N	N	%
3rd attempt	3	2	5	7,1
Pass practical driving test				
1st attempt	32	20	51	74,3
2nd attempt	11	6	17	24,3
3rd attempt	-	1	1	1,4
Number of accidents				
None	27	14	41	58,6
One	12	12	24	34,3
Two	4	1	5	7,1
Three or more	-	-	-	
Thereof accident with				
Only material damage	15	13	27	96,6
Personal injury	1	-	1	3,4

The participants passed their driving test on average 20 months before their participation in FSF (ranging from 6 to 60 months, MD = 16,5 months). Generally they had driven about 30,000 kilometres with a car since their driving test (range from 90 to 200,000 kilometres, MD = 15,000 kilometres).
Most of the FSF participants reported a frequency of "daily driving" (71.4 %), followed by "several times per week" (25.7 %), and "on average once per week" (2.9%).
The most frequently reported driving purpose in the group of the young drivers was "driving to work or school" (41.2 %), "driving for leisure activities" (36.8 %) and "shopping" (8.8 %).
Generally they came to know about FSF at their driving school (78.6 %), or they heard about it from friends (10 %). Very few (less than 3 % for each source) heard about FSF from newspapers, the Internet, parents and relatives, or others.

2.2.2 Sample characteristics seminar leaders (driving instructors)

229 (95.8 %) of the 246 seminar leaders in the sample were male (7 did not specify gender). The average age of all seminar leaders involved was about 48 years (range 28 – 68, MD = 49). They lived in a uniform distribution across the federal states of Germany. Only 12.9 % of them were employed driving instructors, the other 87.1 % were self-employed. On average they had 23 years of experience as driving instructors with a range from 2 to 46 years of professional experience. Only 11 of the 246 seminar leaders had carried out any FSF seminars so far.

2.2.3 Sample characteristics track instructors

95.9 % of the track instructors in the sample were male. The average age of all driving instructors involved was about 46 years with a range from 24 to 68 (MD = 45). The track instructors' place of residence was spread throughout the federal states of Germany.
25 (17.9 %) of them were full-time track instructors; the others were employed in a second job as track instructors. On average they have worked as track instructors for 8 years with a range from 1 to 32 years.
26 of the 146 track instructors reported having carried out an FSF seminar so far.

2.2.4 Sample characteristics trainer of seminar leaders

On average, the trainers of the driving instructors were 50 years old. 91.5 % of all trainers of seminar leaders are male. On average, each of them had conducted 3 training seminars for seminar leaders so far.

2.2.5 Sample characteristics trainer of track instructors

The trainers of the track instructors were on average 47 years old. 80 % (4 out of 5) of them were male. They had conducted 5 training seminars for track instructors on average.

2.2.6 Sample characteristics programme authors, consultants, multipliers

Neither age nor sex was asked for in the group of the programme authors, consultants, and multipliers. We only asked them for other roles they fulfil in the FSF programme.
3 out of 8 also took on the role of trainer for the trainers of the seminar leaders, 2 out of 8 work as a trainer for the track instructor in FSF, 4 performed as trainer for the seminar leaders, 1 as track instructor and 4 as seminar leader in FSF. 6 of the 8 authors / multipliers had more than a dual function in FSF.

2.3 Rate of return

Altogether 970 questionnaires were sent out and in the end 523 questionnaires were returned (return rate of 54 %). Details on the allocation of these questionnaires to the six groups are shown in Table 32. Most of the questionnaires were not completely filled out. The seminar trainer and the track instructors in particular left blanks or gave no answers to several questions, mostly because they have not implemented any courses yet, so they could not answer the specific questions on organisation, available time for programme modules, etc.
We decided to use as much information as possible from the questionnaire so we did not exclude any cases with incomplete questionnaires. That is why the N in the presented results is not equal to the different variables within the groups.
Complete questionnaires were returned only by 32 participants, 30 track instructors, 3 trainers of the track instructors, and 7 authors.

2.4 Selection bias

All people taking part in the FSF training were addressed and a return rate of 54 % was achieved.
We assume that those people who answered the questionnaire are highly motivated concerning the FSF-model. Conclusions from the rate of return towards a selection bias can not be drawn.

3 Training programme

3.1 Background

Road accidents are the main cause of death in the age group of young drivers. Experts agree that the initial period after getting the driving licence is a very dangerous one for novice drivers, irrespective of the kind of driving lessons given previously.

In this period, driving involves a lot of new traffic and social situations which cannot be effectively addressed in the pre-licence phase, e.g. driving in fog or on slippery roads, driving with passengers in leisure time, driving under the influence of emotions, etc. It is impossible to deal with all these factors in pre-licence driving lessons because young people have no experience with these factors yet and most of these situations cannot really be trained. According to the "good driver hierarchical model" (GDE matrix, see Hatakka et al., 1999) these factors are mainly located on the higher levels 3 and 4.

With these guidelines in mind, the German Road Safety Council (DVR) set up a working group in 1996 to design a second-phase education model to assist young drivers with their experience in traffic. Members of this working group were the German Road Safety Council (DVR), the Ministry of Transport, some federal states of Germany, the German Insurance Association (GDV), the German Federation of Driving Instructor Associations (BFV), and the ADAC acting on behalf of German automobile clubs. The Federal Highway Research Institute (BASt) also took part in the working group.

At that time, discussion was carried out in the DVR working group on whether to introduce an obligatory second-phase education for all young drivers in Germany, but there was not enough support for this in practice. As a result, the DVR working group designed the FSF voluntary second-phase model.

The DVR wrote, examined, and optimised the manuals necessary for the training in 2003. The manuals incorporated the results of the EU ADVANCED project. In 2003 the DVR started to train the future seminar leaders and track instructors. At the same time an Internet website was developed to provide information for the target group.

3.2 Description of the FSF-model

3.2.1 Participants

The training is designed for 6 – 12 participants. Every novice driver with a probationary category B driving licence can participate. Participation is voluntary. The participants need to have passed their driving license at least six month beforehand. With their participation they can reduce the probationary time from two years to a minimum of one year (if a novice driver seriously violates traffic regulation during the probation period, the probationary time is extended from two to four years) or from four years to a minimum of three years. Upon participation in FSF, the probation period can only be shortened once. Only those parties who reside in one of the federal states of Germany that has introduced the second-phase training by law are eligible to participate. Participants can register directly at driving schools that offer the seminar.

The DVR is expecting the following four categories of participants:
1. young persons who aim for advanced driving training (intrinsic motivation)
2. young persons who are encouraged by their parents or other relatives to participate (more extrinsic motivation)

3. young persons who have already committed a traffic offence within the probation period and know how quickly you can get involved in dangerous situations. They take part on the one hand to become a safer driver and on the other hand to reduce their probationary time from the extended four years to three years.
4. Young people who consciously want to shorten their probation time from two to one year to evade the specific control, monitoring, and constrictions laid down during the probationary period.

3.2.2 Course of events

The programme is implemented by two trainers: A seminar leader (driving instructor) and a track instructor.
Based on two manuals, the seminar leaders and the track instructors implement the FSF programme. The course of action is strictly structured. The track instructors' manual in particular gives very clear operational instructions.

The programme consists of five different seminar modules taking place on five separate days:
- Three group discussion sessions with a duration of 90 minutes each, conducted by the seminar leader.
- An exercise and observation feedback drive in public traffic with a participant-driver, one or two passengers (other FSF participants) and the seminar leader. The individual driving time per participant is at least 60 minutes.
- Practical safety training on a non-public exercise track with a track instructor (duration 240 minutes).

Normally the programme order alternates between group discussions and the practical driving components. The whole seminar needs to be accomplished within a period of two to eight weeks. Only one seminar part may be carried out per day because the young drivers need sufficient time to reflect on their own experiences in the seminar for developing and implementing new safety strategies when driving. With this in mind, the DVR recommends the utilisation of the total permitted period of eight weeks for the five modules.

The following Table 34 specifies the content of the modules in the five seminar parts.

Table 34: Specific modules of the five FSF programme parts

1st Group discussion	o Getting to know each other o Driving experience and reasons for driving o What do I do differently now from when I was in driving school? o What can happen? – experiences with tricky and dangerous traffic situations o What do I still find difficult and unpleasant? What do I want to practise?
Exercise and observation drive	o Reciprocal observation of normal driving style, based on exchanges in the group of drivers o Practising situations which were individually identified as difficult and unpleasant in the first meeting o Modern driving – environmentally friendly and energy saving driving according to instructions
2nd Group discussion	o "Together on the road" – influence of passengers o "Driving- the nicest triviality in the world?" – divided attention while driving o "Excitement, annoyance, time pressure & other factors" – what influence do emotions have on driving? o Optional "Who is actually guilty?" – sharing responsibility in practice
Practical safety training	o Emergency braking ("How well do I still do it? How can I still improve it? That is optimal!") o Are the ideas of the braking distance correct? (e.g. with increased speed) o Braking on slippery streets o Do I brake alone? (Braking with passengers) o Braking in surprise situations o Appropriate braking o Driving curves at a comfortable speed (safely and relaxed) o Driving curves with passengers o Too fast around curves
3rd Group discussion	o "Preventing driving under the influence of alcohol and drugs" o "The wheel turns full circle" – the diversity of traffic (e.g. how was it in the practical driving sessions?) o "Now I know what to do" – personal strategies for driving without accidents and to master all driving tasks more competently

3.2.3 Methods used

During the seminar, several participant-orientated methods and variants are used, such as moderation, behaviour monitoring, behaviour guidance, individual-, partner- and group-work, group-discussion, feedback-discussion, practical exercises and driving training.

3.2.4 Goals of the second- phase education

The main goal of the FSF programme is to discuss the experiences, problems and difficulties of novice drivers. The risk awareness, hazard perception as well as hazard awareness and hazard prevention of the young drivers should be improved. Yet another goal is to practise safe and responsible driving and to convince participants of the sense of safe driving.

The main goals of the German second-phase programme in the five modules are as follows. During the five parts of the programme the participants should:

1st group discussion

- reflect critically on their previous driving experience;
- know the changes between behaviour learned in driving school and their driving behaviour today and should reflect on the reasons for this change;
- examine their driving and their motives for driving with regard to risky driving;
- name unpleasant and difficult situations in traffic to train these situations in the following exercise and observation drive coached by the seminar leader;

Exercise and observation drive

- get feedback on their driving from the other participants (passengers);
- train adequate behaviour in individually-defined unpleasant or difficult situations;
- know and practise the principles of safety, economical and environmentally friendly driving;
- know the benefits of safety reserves: moderate speed and adequate distance;

2nd group discussion

- reflect on the presence of passengers and their influence on safe driving;
- reflect on the influence of emotions and distraction while driving;
- familiarise oneself with typical risk-situations of young or novice drivers;

Practical safety training

- gain practical experience with problematic situations and develop strategies to minimise risk-situations;
- discover the influence of minimal changes in driving behaviour on safe driving;
- improve their skills to use the break properly;
- know how to drive easily and safely through curves;

3rd group discussion

- develop strategies to avoid driving under the influence of alcohol or drugs;
- know that it is correct and important to drive discreetly and by the rules after the probationary time too;
- know that it is important to reflect on their own habits every time they drive; and
- build up safety strategies for the future based on the outcome of the seminar.

3.2.5 Facilities and trainers needed

The "Voluntary Post-Licence Training Course for Probationary Drivers, FSF" is generally offered by specially trained driving instructors (seminar leaders for FSF) at driving schools. The group sessions take place in driving schools; these sessions are lead by one seminar leader. The participants' observation and practice drive is carried out in public traffic with the seminar leader in the vehicle of the seminar leader.
The practical safety training is carried out on a closed track with a professional track instructor. In the practical safety training the participants drive their own vehicles and / or the vehicle that they usually drive (e.g. parents' car).
The seminar leaders and the track instructors need to communicate and discuss the contents of the "practical safety exercises for novice drivers" – based on the experience from the group discussions and observation drive.

3.2.6 Marketing done on the programme so far

At present there are several flyers in circulation, namely from the DVR, several federal states, and a publishing company. These range of flyers describe the FSF training and are practically the same. The German automobile club ADAC and others also published a flyer on FSF which differs slightly from the three others mentioned above.

The DVR and a sponsor from the automotive industry produced an advertising film called "Die zweite Phase" ("the second phase") that is available on CD or DVD and was delivered at least to all driving instructors who are part of the German Driving Instructor Association. With the

help of this film, FSF can be introduced in driving schools within the scope of regular driving lessons. In addition, an information and communication platform was set up on an Internet page www.zweitephase.de. Here driving instructors can download foil slides and samples for advertising or promotion. The official magazine of the German Federation of Driving Instructor Associations (BFV, Bundesvereinigung der Fahrlehrerverbände) "Die Fahrschule" ("Driving School") has published several articles in 2004 that inform the driving instructors on FSF and on further possibilities of advertising for FSF.

3.2.7 Trainers (selection of trainers, training-of-trainers)

The leader of the FSF seminar – except for the module "practical safety exercises for novice drivers" – is a regular driving instructor. The driving instructors who carry out the group discussions and the observation drive in FSF have passed a special train-the-trainer programme for FSF. This train-the-trainer seminar is led by specially trained instructors (trained by DVR coaches such as the FSF authors, consultants, and multipliers from DVR).

The trainers of the future seminar leaders are introduced in a two-day seminar with a maximum of 16 future trainers-of-trainers. A pre-condition for participation on the training of the trainer is that the future trainer of the seminar leaders needs to have already completed the eight-day seminar for trainer-of-the-trainer for the obligatory driver improvement seminar (ASF) for probationary drivers who have committed (severe) traffic offences.

Current seminar leaders (driving instructors) again are instructed by these trainers. This briefing lasts at least one day, mostly two days, and is held by at least one trainer-of-the-trainer and a maximum of 16 future seminar leaders. Before taking part in the seminar for FSF seminar leaders, the driving instructors need to have already participated successfully in the eight day introduction on ASF seminar leader.

In the FSF seminar-leader training course, those attending are informed about the developmental history and the legal background of the programme. They then work through the whole manual (programme-handbook) and participate or practice with role-play, discussions on how the novice drivers may behave, or take part in small group discussions. All participants need to have accomplished at least one FSF programme part themselves during the introduction seminar.

The leader of the FSF track training module is a special trainer from the DVR post-license courses for cars. These track instructors stage the practical safety exercises on track. They need to have attended a special train-the-trainer seminar as well.

Before executing the FSF training the driving instructors need to have accomplished at least five practical training sessions on safe car-driving within the last year. They also need to be experienced in social and educational work with teenagers and adolescents and be associated with a QM-system in accordance with German DIN EN ISO 9001:2000 (QM standards for DVR car safety training).

A concept for a further training course for the seminar leaders and the track instructors after the basic introductory seminar has not been worked out so far.

3.2.8 Experiences on implementing the training (feedback from organisers, trainers, trainees)

Unfortunately, not enough experience has been accumulated at this stage. Only a few seminars have taken place because there is little demand for the training at the moment. The following statements are some individual opinions we picked up during our survey.
A lot of seminar leaders and track instructors involved in the study complain about the lack of demand for the voluntary training. In their view, more efforts need to be made on advertising for

the event and the incentive for participation should be enhanced. A seminar leader comments that the event was a "dead loss" because he could not find any participants for FSF among his learners despite having done a lot of advertising. Another seminar leader said that the FSF training he implemented was profitable for all of his participants, because they were all highly motivated. On the other hand, he complained about the great deal of work that needed to be done before his first FSF seminar took place. Another seminar leader reported on a participant who was clearly not highly motivated and only took part in the training measure "to buy one year of probation time".
One participant said that she considered the participation fee to be very high, especially for young drivers with little money. She also felt that it was not right that the shortening of the probation period was granted to all participants even though not all of the member of her seminar group showed any understanding of the concept of safe driving, had built up a responsible attitude towards alcohol and drugs, nor were they willing to change their inadequate driving style after the event.

3.3 The FSF-model in the context of other post-licence programmes in Germany

Although the standard of driving training in Germany is high in the pre-license area, there has always been great interest in additional training programmes for novice drivers after they have taken their driving test.
Thus the FSF-model is not an entirely new approach in Germany. It enters into a field where a number of other second-phase concepts already exist or have been piloted in the past.
Second-phase measures can be considered in two ways:
(1) as specific educational interventions in the novice drivers' post-licence learning process, in which the newly-achieved driving skills develop and mature, or
(2) as an additional second part of driver training, which is continued in the post-licensing phase.

Concerning the first perspective, it becomes obvious, that a common second-phase event of less than twenty or even ten training hours is a short-term intervention compared to the 2.5 year period of the post-license learning process during which the accident risk will narrow down to 10 % of the initial risk directly after passing the driving test. Considering second-phase programmes as the second part of a formal training period, it seems appropriate to entitle it a "phase", compared to the duration of the first part of formal instruction which lasts in the range of 30 to 50 hours. As the real process, in which novice drivers should acquire safe driving skills is the essential reference frame, it should always be kept in mind that second-phase training measures are only a short-term intervention in a rather long post-license learning process.

In general, the FSF-model can be considered as functionally equivalent to other second-phase training measures which follow a long-standing primary pedagogical orientation towards influencing the attitudes of the novice drivers in the sense of being critical towards one's own skills and possibilities of control, rather than developing manoeuvring skills.

Characteristic elements of the second-phase training measures in Germany are lessons with group discussion, feedback drives and track training. Beside these main elements there is a variety of other information and learning elements, among them simulation of traffic risks, role plays, or personal information from experts.

Group discussions follow a didactical structure, which shifts the activity and the responsibility for the contents of the debate to the participants. Much efforts is put into further training measures to qualify instructors to follow this teaching style adequately.

Feedback drives are novice drivers' test drives in real traffic accompanied by one to two other novice drivers and a driving instructor. After the observation-drive the driver receives some feed-back from the other novice drivers, but generally not from the driving instructor. Then

another novice takes to the wheel until each has performed his / her feedback-drive. The aim is to enhance critical insight into the individual driving style.

In the track training, which follows detailed quality criteria from the German Road Safety Council (DVR), novice drivers are brought into specific driving situations, where they learn about their limited possibilities to control the vehicle with the aim of convincing the novice drivers that it is better to avoid critical situations than to try to master them.

Table 35 overleaf gives an overview of educational second-phase measures actually practised in Germany. Whilst these programmes do not differ in their basic objective, differences can be found in the specific target groups, the way to address the target group, and in the inventory of elements the measures are composed of.

Table 35: Educational second-phase measures in Germany

programme/model	main features	number of participants per year	remarks
FSF-Modell (Freiwillige Fortbildungsseminare für Inhaber der Fahrerlaubnis auf Probe) (DVR)	- group discussions - track training - feed-back-drive - individual, deliberate participation - duration 10-11 hours - pedagogical staff: further trained driving instructors and track-training-instructors - participation leads to a deduction of 1 year probation time	- in the first year several hundred expected - individual access	- started in the beginning of 2004
„EVA – Ernstnehmende Verkehrsansprache“ (Bavaria)	- group discussions - feed-back-drive - pedagogical staff: further trained novice teachers and driving instructors	- whole age-group in vocational schools	- organised by the state ministry for education - started in 2002
„Alles im Griff“ (DVR)	- classroom work - flexible structure based on modules - duration between 1 hour school-lesson and 2-day-seminar	- 20,000 per year - access via vocational schools	-
„Aufbauseminare für Fahranfänger” (DVR)	- group discussions - feed-back drive - pedagogical staff: further trained driving instructors	- compulsory participation for probationary drivers with certain traffic offences - 70,000 per year (= 7-8 percent of the new drivers)	- didactical approach comparable to FSF-model
Pkw-Sicherheitstraining (DVR; different course providers)	- group discussion - instruction - track-training - duration: 1 day - pedagogical staff: further trained safety-training-instructors	- several thousand - individual participation	- directed to all age groups
„Fahr-/Spartraining“ (DVR)	- instruction - group discussion - exercise drive in real traffic - pedagogical staff: further trained safety-training-instructors	- 5,000 - individual participation	- directed to all age groups
„Aktion junge Fahrer“ (Deutsche Verkehrswacht)	- face-to-face information - exercises - demonstrations, presentations - discussion	- several thousand - individual/collective participation	- mixture of informational and educational elements
„Christopherus-Seminare“ (Landesverkehrswacht Niedersachsen)	- group discussion - group work	- access via vocational schools and business companies	-
„Spar-/Sicherheitstraining“ (Landesverkehrswacht Niedersachsen)	- instruction - group discussion - track training - exercise drives	- individual/collective participation	-
„MOVE-Modellprojekt Fahranfänger“ (Landesverkehrswacht Sachsen-Anhalt)	- group discussion - face-to-face information - exercises - demonstrations	- several thousand - access via schools	-
Fahrsicherheitstraining (ADAC Sachsen-Anhalt)	- group discussion - demonstrations - track-training	- access via schools	- special version of the DCR-Pkw-Sicherheitstraining, offered exclusively to young drivers
„Fit for Drive“ (Landesverkehrswacht, ADAC Schleswig Holstein)	- group discussion - track-training	- access via schools	-

On the German national level, two main second phase measures should be mentioned: the DVR-Safety Training for Category B Drivers (Pkw-Sicherheitstraining) and the post-licence training

courses for novice drivers who have committed severe traffic offences (Driver Improvement Courses). These measures are described in detail in the DAN report (Bartl, 2000; p. 89ff).

The DVR-Safety Training Course is a one-day driver training course that comprises a theoretical and a practical part. It is a voluntary measure offered in general to drivers of passenger cars. Frequently, special advertising campaigns are launched to get young and novice drivers into this training course.

The Driver Improvement (DI) Courses are obligatory for novice drivers who have committed severe driving offences. There are two types of DI-Courses, the first type – run by driving instructors – has been conceived for non-alcohol related offences; the second type – run by traffic psychologists – has been designed for alcohol-related offences. The DI-courses held by driving instructors contain group discussions and a feedback drive. The DI-courses carried out by traffic psychologists only involves group discussions. Every year more than 70,000 persons in Germany have to attend one of these courses, which means that approximately 1 in every 13 novice drivers participates in the DI-Courses for offending novice drivers.

Second-phase programmes are also carried out in schools where pupils in secondary schools are addressed ("Alles im Griff", "Fit for Driving", "EVA", "MOVE"). The great advantage of making contact with the target group directly in schools is that larger numbers of young people can be included in the measure. Organising events outside school or another institution, where the target group is already pre-selected in an organised way, results in only very few people taking part in the measure. This effect was especially experienced in the programme "Aktion junge Fahrer", which was offered to the target group on public places and during leisure time. Taking part in these events requires extra time and additional effort on the part of those who are interested. Experience shows, that not many are willing to overcome these barriers. Another advantage of second-phase activities carried out directly in schools or in connection with schools is that a broader and more differentiated offer of pedagogical activities can be made, resulting in possible synergy-effects from combining the extra second-phase activities with ordinary items from the school curriculum.

There have been several evaluations of second phase measures in Germany. None of them could prove that these activities contribute towards lowing the accident risk. The methodically most demanding evaluation was on the pilot project "Jugend fährt sicher" ("Young people driving safely"), which is a predecessor of the actual FSF-model and was run in 1991 to 1993 in a couple of federal states in Germany. In the evaluation (Schulz, Henning & Chaselon, 1995) which followed a before-after-design with control groups, an examination was carried out to whether the model was able to change the attitudes of novice drivers or not, and – in doing so – to lay the foundations for safer driver behaviour. The results showed that attitudes in the treatment group and in the control group developed without any quantitative difference in the same manner: they got markedly worse. This eroding process may be typical for people having just come out of school or another educational process. After leaving school, people change their behaviour from the way they learned it in school to a rather individual and practically adapted style of behaviour. What the evaluation proved was that the second-phase intervention was not able to stop or limit this change for the worse.

The DVR-Safety Training was evaluated several times. The most important evaluations are from Seidel & Beetz (1978) and Kiegeland (1999). None of these evaluations could prove safety effects in terms of reducing the accident risk. Training effects could only be proved for several knowledge and behavioural items. One of these items was a better performance of the emergency brake, which was not tested inside a real traffic situation and therefore covers only the technical but not the situational aspect of operating the emergency brake.

For alcohol related driver improvement courses (NAFA; "Nachschulung alkoholauffälliger Fahranfänger") an evaluation analysed the assumed lower relapse rate in alcohol-related

offences compared to a control group (Jacobshagen, 1997, 1998). Again, there is no direct proof of effects on accident involvement, and furthermore questions are discussed on the comparability of the control-group which was included from another study (Bartl, 2000).

The details mentioned above point out that the FSF-model in Germany is not in unchartered territory. Irrespective of the reported poor evaluation results for educational second-phase measures up to now, safety training course providers, driving schools, and many experts are convinced that this kind of training can contribute substantially towards improving the road safety behaviour of novice drivers on the road.

With the FSF-model the proponents of second-phase training were successful in finding political majorities to introduce a new pilot project into German law and to start it at the beginning of 2004. The law explicitly states that the safety effects of the model must be evaluated by BASt, the Federal Transport Minister's Highway Research Institute. The proponents of the FSF-model, the German Transport Minister and BASt have agreed on an evaluation concept. Following this concept, a two-stage evaluation over a period of six years is scheduled. The first evaluation part will be a process analysis, which serves to optimise the model and its practical execution. The second stage will be an evaluation of the impact of the model to lower the accident risk of its participants. The first steps of the process evaluation were performed within the frame of the NovEV project and the results are laid down in this report.

4 Evaluation design and timetable

In the German evaluation study a one-time survey was performed that is primarily designed to see if designers, trainers of trainers, seminar leaders, track instructors and the participants have similar opinions on the training itself. As such, we will not be evaluating directly the effects of the training on safe driving behaviour or attitudes, although clues on the outcome and attitude changes of the participants can be drawn from several questions within the questionnaire.
It will, however, provide some indication of whether the training is being implemented and received in the manner in which it was intended. The evaluation will also show what type of young driver is taking part in the training (e.g. safety-conscious individuals, people wanting to reduce their probationary period, people wanting to better their driving skills etc.).

4.1 Data collection methods

Six quasi-identical questionnaires were developed for the six groups involved in FSF. Similar scopes were assessed in the questionnaires for all six groups. Following this design, different views on the programme and its outcome with regard to the programme organisation, the programme implementation and the (intended and perceived) participants' achievements were investigated.

The first part of the questionnaire featured information on personal data like sex, age, or domicile. The second part consisted of questions on the organisation and implementation of the FSF programme itself. This gives an insight into whether the programme was implemented the way it was designed and specified by the authors. In a third part of the questionnaire we asked about the time available for all programme parts (trainer only) and the comparative importance of the programme modules for novice drivers' everyday driving. This was done to provide the opportunity to further optimise the FSF seminar and track training in future. If the analysis shows that the settings of time standards are too limited, the time setting needs to be changed. In addition, if the trainers report varying degrees of importance for several parts of the training course, the emphasis of the seminar might need to be adapted according to the demands of the groups involved. The programme as a whole is ranked in accordance to school marks to get a rating of general personal satisfaction with the training programme.

In the next block we compare the different individual experiences with the training programme between the six groups. We ask about the motivations of the participants to attend FSF and the activities they performed during the programme (from different point of view, such as participants and the trainer). The outcome of the training ("What I know or can do better after the programme") is the focus of the next part of the questionnaire. Here the main goals of the programme (as specified by the programme authors) are rated by all persons from the perspective of the participants. In a final step, the main topics of the training programme are assessed according to the four levels of the GDE matrix. This is to see whether all groups categorise the parts of the training programme on identical levels or if there are different appraisals in theory (authors) and practise (trainers and participants).

All the information mentioned above was collected within the six groups. While the novice drivers report their point of view on the programme, the trainers report their own impressions and the programme-designer gives an assumption on their supposed intentions and defaults.

All questionnaires for the six groups were mailed out at the same time.

4.2 Statistical analysis

Mostly descriptive statistics such as frequencies, comparisons of means and distribution etc. were used in this study. When significant tests were worked out, chi-square-tests and F-tests (ONEWAY-ANOVA) were used. For multiple comparison procedures of group differences we used the Games-Howell test. This test is recommended for the case of assumed heterogeneous variances and unequal sample sizes. The merit of the Games-Howell test is that this procedure is the clear choice if the population variances are known to be unequal and fortunately the Games-Howell procedure does not lead to a substantial loss of power relative to alternative procedures that assume equal variances (Kirk, 1995).

All analyses were conducted with SPSS 10.0.

5. Evaluation Results

In the subsequent tables and figures the following abbreviations for the six groups are used:

PART	Participants
SL	Seminar leader
TI	Track instructors
TR SL	Trainers of the seminar leaders
TR TI	Trainers of the track instructors
MULTI	Authors / multipliers

Conceptually accurate implementation of the training programme
Organisational aspects of the seminar

Questions on organisational topics of the seminar were addressed to the participants as well as to the seminar leaders to see if the concept is being implemented the way it was intended to by the authors. Here we asked questions about procedures with worksheets (given to and filled out by participants during training) or how the seminar leaders dealt with the reminder that was specified in the manual to be sent out to the participants three months after the FSF event.

Table 36 gives an overview of the answers regarding these topics in the groups of participants and seminar leaders.

Table 36: Organisational aspects of the seminar, perceived by participants and seminar leaders

Item		PART [7]	SL	Chi^2	Sign.
Worksheets were distributed at registration.	N	66	9	0,70	0,32
	Yes %	40,9	55,6		
	No %	59,1	44,4		
Worksheets (homework) were filled out at home.	N	70	9	0,65	0,35
	Yes %	64,3	77,8		
	No %	35,7	22,2		
Worksheets (homework) were discussed at the following seminar.	N	69	9	0,55	0,61
	Yes %	94,2	100		
	No %	5,8	-		
Participants received a folder for the worksheets.	N	69	9	3,42	0,07
	Yes %	34,8	66,7		
	No %	65,2	33,3		
A letter was sent to the participants three months after the event.	N	53	6	0,84	0,66
	Yes %	32,1	50		
	No %	66,0 [8]	50		

What is notable is the significant percentage of participants who did not do their homework ("worksheets were filled out at home"). About 1/3 of the participants reported they did not do their homework (i.e. complete their worksheets). In addition 1/3 of the participants did not receive the compulsory reminder three months after the FSF event. Further attention needs to be given to this by the seminar leaders.

As only a few seminars have taken place so far, only nine seminar leaders (who had already conducted a seminar) were able to report on their experiences with FSF organisation.
The organisational procedure is more or less equally reported in the groups of participants and seminar leaders. The only palpable difference can be found in the item "participants received folder". Here two thirds of the seminar leader confirmed that they handed out a folder, while only one third of the attendees confirmed that they had received a folder. Unfortunately an assignment of participants to their respective seminar leaders was impossible since the data was collected anonymously.

Chi^2 tests showed no significant differences between any of the statements of the participants and seminar leaders.

Time allocation and management

To estimate and evaluate the required time for the seminar modules, the seminar leaders and track instructors were asked if the time to implement the seminar or track modules was sufficient. The main purpose of these questions was to help optimise the time allocations for FSF.

In the seminar

The seminar leaders were asked if the available time for the implementation of the seminar modules was adequate. Possible answers were: too short (coded with "1"), adequate ("2"), or too long ("3"). Again answers were available only from nine seminar leaders.
The findings show that the seminar leaders reported more than necessary time (X = 2.33, SD = 0.5) for the exercise and observation drive only. By contrast, the available time for the group

[7] Abbreviations used here and in all tables below: PART = participants, SL = seminar leaders, TI = track instructors, TR SL = trainers of the seminar leaders, TR TI = trainers of the track instructors, MULTI = authors / multipliers.
[8] %-sum unequal 100 because several training took place less than three months before our survey.

meetings is considered to be too short: The average estimation for the first group meeting is X = 1.67 (SD = 0.5), for the third group meeting X = 1.63 (SD = 0.53) and for the second group meeting X = 1.56 (SD = 0.53). These results indicate that time allocations, especially for the second group meeting, should be revised.
The reported time estimations may not be valid since all seminar leaders who gave their time appraisals were inexperienced in implementing the programme and accordingly their judgements are based on the first implementation of a FSF seminar alone. From our point of view more experience with FSF is needed in order to justify a revision of the time allocations based on seminar leaders' feedback.

On the track

Similar to the seminar leaders, the track instructors were asked if the given time was adequate to accomplish the practical safety training, as specified in the manual. Data is available from 21 track instructors who stated that they had already implemented FSF driver training.
The track instructors reported adequate time standards for the theoretical parts of the track training (X = 2.1, SD = 0.44). In contrast, more time should be available for the practical part of the track training (X = 1.57, SD = 0.51); particularly short sessions were reported more for the "braking part" (X = 1.48, SD = 0.51) than the "driving curves part" (X = 1.90, SD = 0.70) of the FSF track training.

Importance of seminar modules for everyday driving

In the seminar

The participants, the seminar leaders, the trainers of the seminar leaders and the programme authors were asked to assess the importance of the different seminar modules dealt with in the group discussion with regard to everyday driving and fostering safe driving amongst novice drivers.
The importance was rated on a scale from "1" (very unimportant) to "6" (very important).

Table 37 shows the average ratings and standard deviations of importance as rated by the four different groups for each of the seminar modules. For a consistent N in the groups we considered only those questionnaires that rated the importance of all relevant FSF modules. These hypothetical assessments are independent from having implemented the FSF programme; therefore all available ratings are considered in the following table, regardless of whether the seminar leaders have implemented a FSF seminar so far or not.
The five statements considered to be the most important for future safe driving per group are printed in bold. An ANOVA was conducted to analyse differences in the assessments between the groups. FSF elements with significant variation in assessed importance between the groups are shaded grey in the following table.

Table 37: Evaluation of importance of the seminar elements in four participating groups (1 = very unimportant, 6 = very important), five most important elements per group printed bold

Module	PART (N = 44)		SL (N = 44)		TR SL (N = 39)		MULTI (N = 7)	
	X	SD	X	SD	X	SD	X	SD
Group discussions overall [9]	4,70	1,02	5,18	1,08	5,38	0,81	**5,71**	0,49
Exercise & observation drive overall	5,18	1,19	5,25	0,87	5,05	0,92	5,86	0,38
FSF overall	5,20	0,73	5,27	1,09	5,38	0,81	**5,71**	0,49
Group discussions								
Changes in driving behaviour since exam [10]	4,82	0,97	5,25	1,04	**5,46**	0,68	5,57	0,35
Experience with difficult situations	**5,27**	0,69	**5,41**	0,97	**5,51**	0,64	5,57	0,79
What I still find difficult	4,75	1,26	5,05	1,16	5,13	0,92	5,43	0,79
Impact of passengers [11]	5,00	1,03	5,23	0,99	**5,51**	0,72	**6,00**	0,00
Distractions while driving	**5,32**	0,80	**5,43**	0,93	**5,46**	0,88	5,57	0,53
Impact of feelings [12]	**5,07**	0,93	**5,50**	0,90	**5,62**	0,63	**5,86**	0,38
Avoidance of alcohol and drugs and driving [13]	**5,57**	0,70	**5,43**	0,95	4,87	0,98	5,29	0,95
Review of topics and results of the seminar	4,64	1,04	4,77	1,05	4,87	0,89	5,43	0,53
Personal strategies for safe driving [14]	4,93	1,15	5,20	1,05	**5,56**	0,64	**6,00**	0,00
Observation Drive								
Observe driving of others	4,93	1,04	5,05	0,96	5,00	0,86	5,57	0,53
Being observed by others	4,95	1,24	5,23	1,01	5,23	0,84	**5,86**	0,38
Analysis and discussion of observation drive	**5,18**	1,21	**5,39**	0,84	5,26	0,85	**5,71**	0,49
Practise difficult situations	5,02	1,27	5,05	0,96	4,97	0,87	5,57	0,79
Environmentally friendly and energy-saving driving	4,95	1,46	4,84	1,18	4,92	0,96	5,57	0,79

Statistical differences were tested with an ANOVA. If the ANOVA indicated significant differences between groups, multiple mean comparisons were conducted with the help of the Games-Howell test. This method is suitable for paired comparisons when the homogeneity of variance cannot be assumed.

The tested interaction effect "group by discrete estimated importance variables" pointed out the following significant differences (F-test, $p < .05$) between groups. Differences were found for the estimated importance of "group discussions overall" ($F_{(3,130)} = 4.59$, $p < .01$), "changes in driving behaviour" ($F_{(3,130)} = 4.14$, $p < .01$), "impact of passengers" ($F_{(3,130)} = 3.76$, $p < .05$), "impact of feelings" ($F_{(3,130)} = 4.21$, $p < .01$), "alcohol and drugs" ($F_{(3,130)} = 4.75$, $p < .01$), and "personal strategies" ($F_{(3,130)} = 4.46$, $p < .01$).

[9] ($F_{(3, 130)} = 4,59$, $p < .01$)
[10] ($F_{(3, 130)} = 4,14$, $p < .01$)
[11] ($F_{(3, 130)} = 3,76$, $p < .05$)
[12] ($F_{(3, 130)} = 4,21$, $p < .01$)
[13] ($F_{(3, 130)} = 4,75$, $p < .01$)
[14] ($F_{(3, 130)} = 4,46$, $p < .01$)

Further analyses of group differences for these variables proceeded. Table 38 shows the particular paired groups with significantly different means.

Table 38: Multiple analyses of group differences in estimated importance of FSF modules, results of Games-Howell test

Module	Sign. differences between group means
Group discussions overall	PART < TR SL ($p < .01$), PART < MULTI ($p < .01$)
Changes in driving behaviour since exam	PART < TR SL ($p < .01$), PART < MULTI ($p < .05$)
Impact of passengers	PART < MULTI ($p < .01$), SL < MULTI ($p < .01$), TR SL < MULTI ($p < .01$)
Impact of feelings	PART < TR SL ($p < .05$), PART < MULTI ($p < .01$)
Avoidance of alcohol and drugs and driving	PART > TR SL ($p < .01$), SL > TR SL ($p < .05$)
Personal strategies for safe driving	PART < TR SL ($p < .05$), PART < MULTI ($p < .01$), SL < MULTI ($p < .01$), TR SL < MULTI ($p < .01$)

The importance of the seminar elements overall is estimated to be very high. Values between 4.7 and 6.0 on a scale from 1 to 6 are reported. As shown in Table 37, there is a certain consensus in the four groups. The participants' and the seminar leaders' five most important seminar elements are quite similar. The "impact of feelings" is rated in the "top 5" of the important seminar parts in all four groups. Other seminar sections are in the "top 5" of at least three groups: these are "experience with difficult situations", "distraction while driving", and "discussion of observation drive".
While the authors and the trainer of the seminar leader regard "personal strategies for safe driving" as one of the most important topics in FSF, the seminar leaders and participants judge this topic to be of secondary importance in FSF. More emphasis on the importance of this topic needs to be placed in further training of the seminar leaders; in particular because the differences in the assessments of these groups are significant (see Table 38).

On the track

In similar fashion, the importance of the track exercises for future everyday safe driving was assessed by the participants, the track instructors, the trainers of the track instructors and the authors of the programme. On average the participants reported higher importance for the track training parts then for the group session parts.
Table 39 gives an overview of the importance given to the track training parts in the different groups. Items with significant differences between the groups are shaded grey. The five most important FSF track parts for novice drivers' everyday driving in every group are printed in bold.

Table 39: Estimated importance of the practical safety training (PST) elements in four participating groups (1 = very unimportant, 6 = very important), five most important elements per group printed bold

Module	PART (N = 44)		TI (N = 88)		TR TI (N = 5)		MULTI (N = 7)	
	X	SD	X	SD	X	SD	X	SD
Theoretical part PST	5,00	0,94	4,58	1,18	**5,40**	0,89	5,29	0,76
Practical part PST	5,59	0,84	**5,40**	0,89	**5,80**	0,45	**5,71**	0,49
FSF overall	5,20	0,73	5,30	0,97	**5,60**	0,55	**5,71**	0,49
Introduction braking [15]	5,64	0,61	5,15	1,02	4,40	0,55	5,29	0,49
Emergency braking without instruction [16]	5,57	0,85	4,92	1,25	**5,40**	0,55	**5,71**	0,49
Preparation for optimal emergency braking	5,68	0,56	**5,37**	0,95	**5,40**	0,55	5,43	0,79
Practice emergency braking	**5,91**	0,29	**5,55**	0,96	**5,60**	0,55	5,57	0,53
Emergency braking with passengers [17]	5,64	0,81	4,99	1,39	5,00	1,22	**5,71**	0,76
Estimation of braking distance	**5,80**	0,46	**5,44**	1,00	**5,80**	0,45	**5,86**	0,38
Braking on slippery roads [18]	**5,91**	0,29	**5,45**	1,03	**5,80**	0,45	**5,86**	0,38
Emergency braking with surprises [19]	**5,73**	0,85	5,15	1,25	**5,40**	0,89	**5,71**	0,49
Outline primary findings	**5,75**	0,49	**5,44**	0,97	**5,40**	0,89	**5,71**	0,49
Introduction curves	5,39	0,75	4,99	1,02	4,80	0,84	5,00	1,00
Driving curves with comfortable speed	5,18	1,13	4,76	1,23	5,00	0,71	5,43	0,79
Driving curves with passengers with comfortable speed	5,30	1,02	4,77	1,35	5,20	1,10	5,29	1,11
Driving curves faster than comfortable speed	5,64	0,87	5,28	1,17	**5,40**	0,89	5,57	0,53
Outline primary findings	5,66	0,53	**5,37**	1,00	**5,40**	0,89	5,29	0,76

Consensus amongst the different groups can be found especially in the assessments for "estimation of braking distance", "braking on slippery road", and "outline primary findings of the topic braking". All four groups regard these FSF parts as one of the five most important in the practical safety training.

Significant differences between the groups appeared for "introduction braking" ($F_{(3, 140)} = 4.67$, $p < .01$), "emergency braking without instruction" ($F_{(3, 140)} = 4.06$, $p < .01$), "emergency braking with passengers" ($F_{(3, 140)} = 3.26$, $p < .05$), "braking on slippery road" ($F_{(3, 140)} = 3.19$, $p < .05$), and "emergency braking with surprises" ($F_{(3, 140)} = 2.95$, $p < .05$).

To further analyse the differences between the groups for these variables, the Games-Howell test was carried out. Table 40 shows the significant differences between the specified groups regarding the importance of the FSF programme parts.

[15] ($F_{(3, 140)} = 4,67$, $p < .01$)
[16] ($F_{(3, 140)} = 4,06$, $p < .01$)
[17] ($F_{(3, 140)} = 3,26$, $p < .05$)
[18] ($F_{(3, 140)} = 3,19$, $p < .05$)
[19] ($F_{(3, 140)} = 2,95$, $p < .05$)

Table 40: Multiple analyses of group differences in estimated importance of FSF modules, results of Games-Howell test

Module	Sign. differences between group means
Introduction braking	PART > TI ($p < .01$), PART > TR TI ($p < .05$)
Emergency braking without instruction	PART > TI ($p < .01$), TI < MULTI ($p < .05$)
Emergency braking with passengers	PART > TI ($p < .01$)
Braking on slippery roads	PART > TI ($p < .01$)
Emergency braking with surprises	PART > TI ($p < .05$)

Judging the event with the help of school marks

The FSF participants evaluated the seminar overall with the help of school marks. This was done to ascertain the participants' personal assessment of the FSF programme.
The German school marks range from 1 ("very good") to 6 ("insufficient"). The most frequent appraisal of the participants was a "2" (55.1 %), which means "good", followed by "1" ("very good", 29.0 %). 11.6 % rated FSF with "3" ("satisfying"), 2.9 % with "4" ("sufficient"), and one participant (1.4%) rated the programme with "5" ("deficient"). Altogether the novice drivers were quite content with the FSF programme: more than 84 % of them rated the programme "good" or "very good".

Individual-benefit assessment

The FSF participants were asked how much they had had to pay for the FSF training course and how many euro they personally think it was worth. On average the participants paid € 169.21 for the training courses with a range from € 0 to € 280; the median is MD = € 190 and the mode is MO = € 200.
Results indicate that they are willing to pay € 169.49 on average with a range from € 25 to € 1000 (!). We expect the € 1000 declaration to be a joke; without this case the average figure of the "value estimation" is € 157.09 with a range from € 25 to € 300. The average difference between the price actually paid and the price considered justifiable by participants is € 7.90 with a range ± € 150. On average, the participants are willing to pay € 7.90 less than they actually paid for FSF participation.
20.9 % report FSF was worth exactly the price they had paid; 34.4 % are willing to pay more and 44.8 % would have paid less for FSF as measured according to their personal benefit.
The participants could estimate the personal cost satisfaction rate in five steps from "1" ("very bad") to "5" ("very good"). On average they reported a satisfaction of X = 3.65 (SD = 1.02; MD = 4) which means a rather positive response to the question of whether they are satisfied with the cost-benefit equation of the FSF programme. 52.9 % chose "good" or "very good", 34.8 % "appropriate" and 11.8 % reported a "bad" or "very bad" personal cost satisfaction level.

The seminar leaders and the track instructors should estimate the cost-benefit relation of FSF from the novice drivers view as well. The majority of the seminar leaders (48 %) considered the cost-benefit ratio "adequate" (3) on a scale from (1) "very bad" to (5) "very good". 27.4 % rated the cost benefit worse than "adequate", and 24.6 % rated better. The track instructors also mainly rated the cost benefit as adequate (42.1 %), 35.6 % rated it worse, and 22.3 % rated it better than "adequate".

The seminar leaders and the track instructors were also asked how many euro they had spent for their own instruction for the FSF training or for their training of the trainer and what they

thought would be a reasonable price for the training. Table 41 reports the cost-benefit responses of the participants as well as of their trainers.

Table 41: Individual benefit estimations of the FSF training and training of the trainer from the viewpoint of the participants and the various trainers

€	PART (N = 67)				SL (N = 148)				TI (N = 75)			
	X	SD	Min	Max	X	SD	Min	Max	X	SD	Min	Max
€ actually spent on training	169,21	73,25	0	280	155,25	97,76	0	600	136,95	160,19	0	1000
Estimation of how much training is worth	157,09	120,20	25	300 (1000)	133,15	80,82	0	450	136,26	114,50	0	500
Average difference between amount spent and estimation	7,90	69,78	-150	150	22,29	75,90	-240	300	11,22	136,11	-250	500

On average, the seminar leaders spent € 155.25 for their training. 42.7 % think that the price they paid is adequate, 13.5 % would have paid more, and 43.8 % reported that they have paid more than they think it was worth.
The track instructors paid an average of € 136.95. 35.7 % of them are satisfied with the price they paid, 35.7 % of them are willing to pay more and 28.6 % of the track instructors think they paid more for their training than it was worth.

Participants' motivation

Participants were asked to state why they chose to participate in FSF. The other persons involved in the FSF programme were also asked to give their opinion on the main reasons for novice drivers for taking part in FSF.
The frequency of answers in the different groups is shown in Table 42; the most frequently named statement in each group is printed in bold.

Table 42: Main reasons for participation in FSF from different views (% in column, mode per group printed in bold)

%	PART (N = 64)	SL (N = 159)	TI (N = 113)	TR SL (N = 47)	TR TI (N = 5)	MULTI (N = 8)
shorten probation period from 2 to 1 year	**26,12**	34,38	31,90	28,26	**30**	**31,25**
shorten probation period from 4 to 3 years	23,13	**35,80**	**32,76**	**32,61**	20	**31,25**
learn more about safe driving	17,16	2,84	2,59	5,43	10	6,25
improve driving skills	22,39	5,68	2,59	1,09	10	6,25
exchange driving experience	2,24	1,70	0,86	2,17	0	6,25
invited by the organiser	3,73	2,84	3,88	5,43	0	0
participation was sponsored	2,24	9,66	16,38	14,13	10	6,25
parents (or relatives) considered it important	2,24	6,25	8,62	10,87	20	12,5
other reasons	0,75	0,85	0,43	0	0	0

Clearly, the main motivation for FSF participation is to shorten the probation period from 2 to 1 or from 4 to 3 years. All six groups agree on that. Another important reason for novice drivers' participation is the wish to improve their driving skills or learn more about safe driving. 22.4 % and 17.16 % of the participants respectively mention these topics while only a small proportion of the trainers, trainers of the trainers or programme authors expect the novice driver to take part in FSF for these reasons. On the other hand, the seminar leaders, the track instructors, and the

trainer of the track instructors in particular feel that the novice drivers take part because they were sponsored or because parents stress the importance of participating.

Participants' activities during the training (intensity of participants' attendance)

The training participants were asked about their different activities during the FSF programme. The activities given in the questionnaire were taken from the programme description in the FSF manual. We also added two "dummy" variables of activities that have nothing to do with the FSF programme; these were "learned something about traffic rules" and "learned to reverse into a parking space".
The following results are based on N = 69 participants. They rated their activity on a scale from "1" ("does not apply at all") to "6" ("applies fully").
The most popular activity among the participants in the group discussions was "reporting on driving experience", followed by "talking about passengers' influence" and "anticipatory driving". The most popular activity in the track training was "differences in driving different vehicles" and "braking distance at different speed" followed by "differences in others driving behaviour". Table 43 (group sessions) and Table 44 (track training) show the activities in FSF from the participants', seminar leaders', track instructors' and authors' point of view. The five most frequently named activities per group are printed in bold.
Two ANOVAs were conducted to check group differences: one for the group discussion part (PART, SL, MULTI) and one for the track training part (PART, TI, MULTI). Items with significant mean differences between the groups are shaded grey in Table 43 and Table 44.

Table 43: Overview of the reported activities (participants) and the expected activities (seminar leaders and authors) in the group sessions, five most frequent activities per group printed in bold

Module	PART (N = 69)		SL (N = 152)		MULTI (N = 8)	
	X	SD	X	SD	X	SD
In the group sessions I						
reported on my own driving experience	**5,30**	1,06	**5,41**	0,71	**5,63**	0,52
talked about common dangerous situations	4,80	1,28	**5,04**	1,00	4,63	1,69
talked about typical specific dangerous situations for novice driver [20]	4,55	1,31	**5,03**	1,03	5,00	1,07
talked about anticipatory driving	4,52	1,36	4,71	1,10	4,63	1,19
talked about predictable behaviour of the other road users [21]	4,29	1,56	4,82	1,12	4,88	0,83
talked about the effects of emotions on driving [22]	4,54	1,33	4,99	1,02	**5,38**	0,74
talked about environmental influences on driving [23]	3,67	1,71	4,43	1,22	4,75	1,28
talked about the effect of alcohol and drugs on driving behaviour	**4,91**	1,31	4,96	1,12	4,88	1,36
discussed the influence of passengers on driving behaviour	**5,03**	1,12	**5,32**	4,66	**5,25**	1,04
learned something about traffic rules	3,68	1,65	3,38	1,53	2,75	0,89
developed intentions and behavioural strategies for safe driving [24]	4,46	1,32	5,02	1,10	**5,63**	0,52
learned something about the necessity of anticipatory driving at different speeds	**4,96**	1,10	**5,04**	1,05	**5,25**	0,71
learned something about the necessity of keeping a safe distance	**4,84**	1,23	4,86	1,00	**5,25**	0,71
learned something about numerous grades of danger in miscellaneous traffic situations	4,75	1,21	4,96	1,00	4,75	0,71

[20] ($F_{(2, 226)} = 4{,}30$, $p < .05$)
[21] ($F_{(2, 226)} = 4{,}35$, $p < .05$)
[22] ($F_{(2, 226)} = 4{,}71$, $p < .05$)
[23] ($F_{(2, 226)} = 7{,}75$, $p < .01$)
[24] ($F_{(2, 226)} = 7{,}19$, $p < .01$)

Table 44: Overview of the reported activities (participants) and the expected activities (track trainers and authors) during the track training, five most frequent activities per group printed in bold

Module	PART (N = 69)		TI (N = 121)		MULTI (N = 8)	
	X	SD	X	SD	X	SD
At the track training I						
meet with problematic driving situations	4,62	1,65	4,75	1,25	5,13	1,13
learned something about the effect of changing driving conditions	**5,10**	1,23	**5,27**	0,83	**5,50**	0,76
examined my self-assessment [25]	4,86	1,17	**5,25**	1,04	**5,63**	0,52
learned something about the effects of road conditions	5,06	1,32	5,09	0,88	**5,63**	0,74
learned something about distraction and driving [26]	4,07	1,64	**5,20**	1,01	**5,63**	0,52
recognised differences in the driving behaviour of other participants [27]	**5,28**	0,95	4,40	1,28	4,88	1,36
recognised differences in the driving behaviour of different vehicles [28]	**5,29**	1,13	4,27	1,38	3,88	1,64
recognised the different braking distance from different speeds	**5,29**	1,04	**5,42**	1,01	**5,63**	0,52
learned something about seating position and safe driving	**5,13**	1,22	**5,10**	1,29	5,13	1,13
learned to back into a parking space	2,00	1,40	1,64	1,01	1,25	0,71

The following Table 45 shows the significant group differences in multiple group comparisons and the direction of the differences.

Table 45: Multiple analyses of group differences in assessment of participants' activities in FSF, results of Games-Howell test

Module	Sign. differences between group means
talked about typical specific dangerous situations for novice drivers	PART < MULTI ($p < .05$)
talked about predictable behaviour of the other road users	PART < MULTI ($p < .05$)
talked about the effects of emotions on driving	PART < SL ($p < .05$), PART < MULTI ($p < .05$)
talked about environmental influences on driving	PART < SL ($p < .01$)
developed intentions and behavioural strategies for safe driving	PART < SL ($p < .01$), PART < MULTI ($p < .01$), SL < MULTI ($p < .01$)
examined my self-assessment	PART < MULTI ($p < .05$)
learned something about distraction and driving	PART < TI ($p < .01$) PART < MULTI ($p < .01$)
recognised differences in the driving behaviour of other participants	PART > TI ($p < .01$)
recognised differences in the driving behaviour of different vehicles	PART > TI ($p < .01$)

[25] ($F_{(2, 195)} = 3,84$, $p < .05$)
[26] ($F_{(2, 195)} = 19,47$, $p < .01$)
[27] ($F_{(2, 195)} = 12,10$, $p < .01$)
[28] ($F_{(2, 195)} = 14,63$, $p < .01$)

The activity profiles are further clarified in Figure 15 and Figure 16. Altogether the statements on activities during the programme are quite similar between the groups. The trends of the activities are almost parallel and no outliers are obvious on first sight. Just as expected, the two "dummy" variables we added were rated with the lowest activity scores.
In the group discussion, the authors of the FSF programme expected higher ratings than the participants actually reported especially in the modules "predictable behaviour of other road users", "effects of emotions", environmental influences" and most important on "developed behavioural strategies for safe driving".

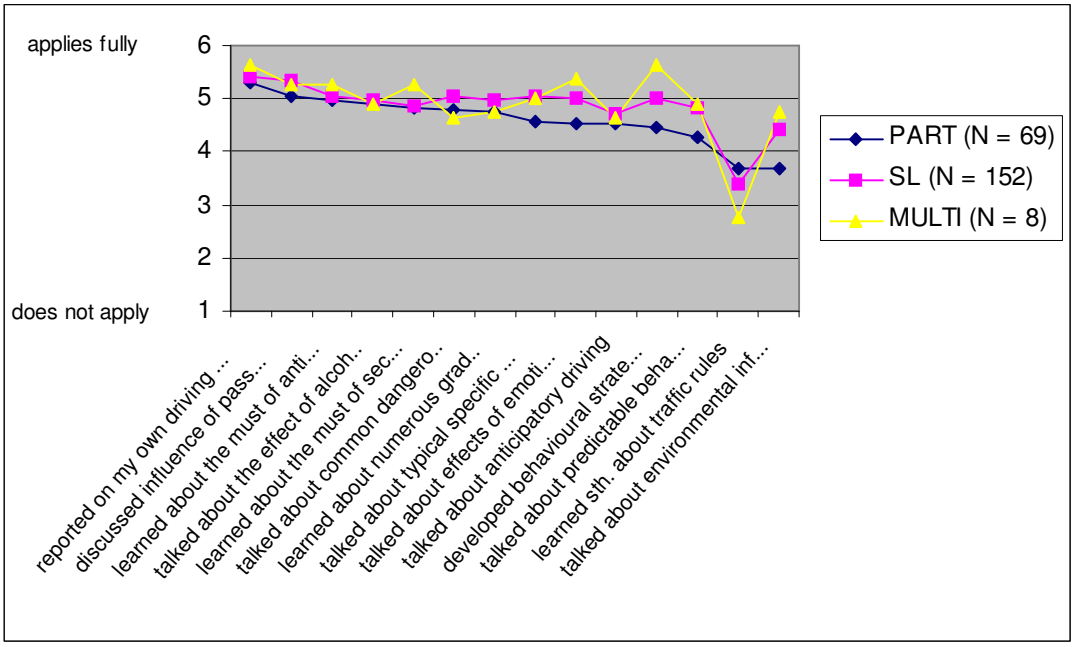


Figure 15: Profile of the estimated FSF group sessions activities in the three groups, sorted by participants' estimation magnitude

In the track training the programme authors expected higher ratings with regard to the participants' critical revision of their self-assessment and in the training part where the participants were supposed to learn something about distraction while driving. The participants gave better ratings than the authors expected with regard to recognising differences in the driving behaviour of other participants and in the driving behaviour of different vehicles.

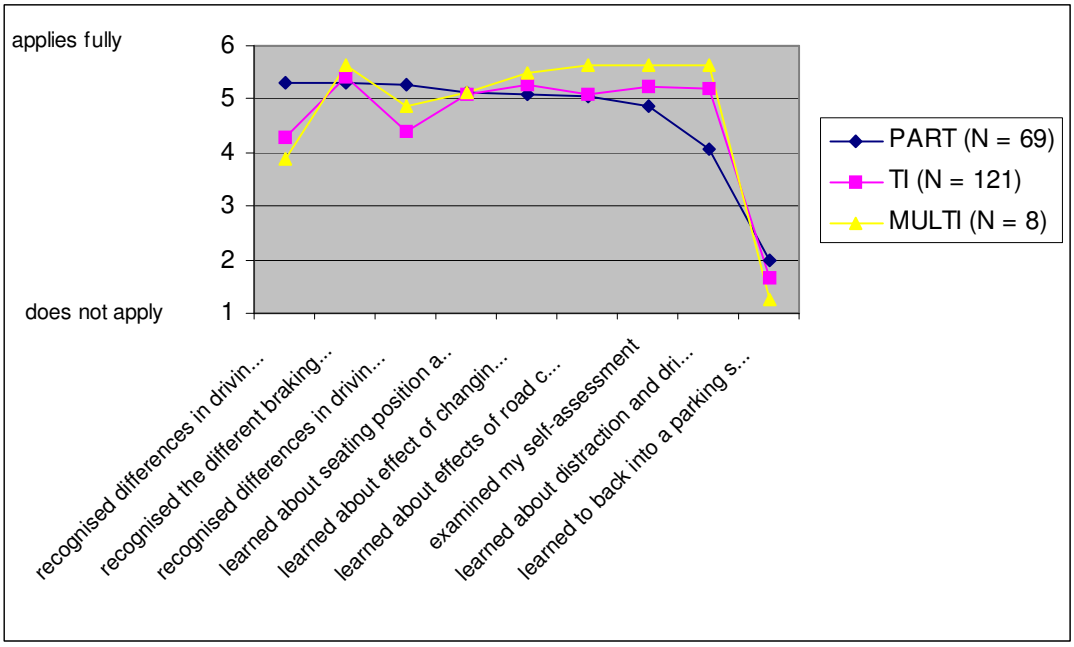


Figure 16: Profile of the assessment of FSF track training activities in the three groups, sorted by participants' estimation magnitude

Positive outcomes for participants

A primary objective of the FSF training was the creation of strategies for safe driving among the novice drivers. We asked the participants how many of these strategies they developed during the event and how often they had been able to implement these strategies in everyday practice ever since.
7.4 % of the participants reported no development of any strategy, 8.8 % reported one, 33.8 % two, 16.2 % three, and 33.8 % four strategies or more.
In road traffic they could apply these strategies not at all (14.7 %), once per month (8.8 %), once per week (5.9 %), several times per week (41.2 %), and several times per day (29.4 %).
Table 46 summarises the results on the strategies developed. E.g. 12 participants reported they developed four or more strategies and use them several times per day; another 15 participants reported they developed two strategies and use them several times per week etc.

Table 46: Development of strategies during FSF and use of these strategies in participants' everyday driving (N = 68)

N	make use of strategies					
amount of strategies developed	not at all	once per month	once per week	several times per week	several times per day	Total
none	5					5 (7,4%)
one	2	1	1		2	6 (8,8%)
two	2	2	1	15	3	23 (33,8%)
three		1	1	6	3	11 (16,2%)
four or more	1	2	1	7	12	23 (33,8%)
total	10 (14,7%)	6 (8,8%)	4 (5,9%)	28 (41,2 %)	20 (29,4%)	68 (100%)

We also asked the participants what they know or can do better regarding car driving after their training. In the questionnaire they needed to appraise twenty statements derived from the FSF

programme goals named in the manual (and one "dummy": "know traffic rules") with the predetermined categories "no changes" (1), "slightly better (2)", and "much better" (3).
The greatest average improvement that participants reported was "braking abilities", followed by "estimation of ones own strengths and weaknesses".
The other groups (seminar leaders, track instructors, respective trainers of the trainer, and authors) were also asked for their assessment of positive outcomes for participants as a result of the training.
Table 47 gives an overview of the average participants' outcomes following FSF in all six groups. The average values of the five statements with the greatest reported (participants) or estimated (other groups) improvement per group are printed in bold. An ANOVA showed significant differences between the improvement reports in the six groups for several items; these items are shaded grey. To go into detail on multiple differences between the groups, we analysed multiple group differences with the help of Games-Howell test. Significant differences between the respective groups are also shown in the left column of Table 47.

Table 47: Estimated participants' outcome ("know or can do better") in the six groups (mean and standard deviation), results of multiple comparisons between the groups; five most improved effects per group printed bold

What I know or can do better after FSF:		N	X	SD
Effects of vehicle characteristics on safe driving [29]	PART	67	**2,16**	0,73
	SL	158	**2,19**	0,53
-	TI	113	2,09	0,51
	TR SL	45	2,07	0,45
	TR TI	5	1,40	0,55
	MULTI	8	1,88	0,64
	total	396	2,13	0,56
Safe control of the vehicle [30]	PART	67	**2,16**	0,69
	SL	158	1,99	0,54
PART > TI (p < .01)	TI	113	1,80	0,55
	TR SL	45	1,91	0,42
	TR TI	5	1,40	0,55
	MULTI	8	2,00	0,53
	total	396	1,95	0,57
Know traffic rules	PART	67	1,31	0,61
	SL	158	1,34	0,55
	TI	113	1,17	0,42
	TR SL	45	1,22	0,42
	TR TI	5	1,00	0,00
	MULTI	8	1,25	0,46
	total	396	1,27	0,51
Estimate influence of driving purpose on safe driving [31]	PART	67	1,93	0,72
	SL	158	2,04	0,60
SL > TI (p < .01)	TI	113	1,73	0,55
TI < TR SL (p < .01)	TR SL	45	2,16	0,56
	TR TI	5	2,00	0,71
	MULTI	8	1,88	0,64
	total	396	1,94	0,62
Assess specific sequence of the planned trip [32]	PART	67	1,54	0,68
	SL	158	1,71	0,61

[29] ($F_{(5, 390)} = 2,70$, $p < .05$)
[30] ($F_{(5, 390)} = 4,82$, $p < .01$)
[31] ($F_{(5, 390)} = 4,81$, $p < .01$)
[32] ($F_{(5, 390)} = 5,59$, $p < .01$)

What I know or can do better after FSF:		N	X	SD
SL > TI (p < .01) TI < TR SL (p < .01)	TI	113	1,37	0,54
	TR SL	45	1,82	0,65
	TR TI	5	1,40	0,55
	MULTI	8	1,63	0,74
	total	396	1,59	0,63
Estimate the effect of the destination on safe driving [33]	PART	67	1,42	0,63
	SL	158	1,80	0,67
PART < SL (p < .01) PART < TR SL (p < .05) SL > TI (p < .01) TI < TR SL (p < .05)	TI	113	1,46	0,57
	TR SL	45	1,76	0,57
	TR TI	5	1,40	0,55
	MULTI	8	1,50	0,53
	total	396	1,62	0,64
Travel planning and route selection [34]	PART	67	1,28	0,55
	SL	158	1,60	0,64
PART < SL (p < .01) PART < TR SL (p < .05) SL > TI (p < .01) TI < TR SL (p < .01)	TI	113	1,28	0,49
	TR SL	45	1,64	0,61
	TR TI	5	1,40	0,55
	MULTI	8	1,50	0,53
	total	396	1,46	0,60
Estimate necessary travel time [35]	PART	67	1,39	0,67
	SL	158	1,66	0,61
SL > TI (p < .05)	TI	113	1,46	0,55
	TR SL	45	1,64	0,57
	TR TI	5	1,20	0,45
	MULTI	8	1,50	0,53
	total	396	1,55	0,61
Response and reaction to passengers' unsafe behaviour [36]	PART	67	1,90	0,76
	SL	158	**2,19**	0,63
PART < TI (p < .05) PART < TR SL (p < .01)	TI	113	**2,34**	1,08
	TR SL	45	**2,44**	0,50
	TR TI	5	**2,20**	0,45
	MULTI	8	**2,38**	0,52
	total	396	2,21	0,81
Self-control in traffic	PART	67	2,13	0,74
	SL	158	2,13	0,61
	TI	113	2,14	0,48
	TR SL	45	**2,31**	0,51
	TR TI	5	**2,20**	0,45
	MULTI	8	**2,38**	0,52
	total	396	2,16	0,59
Avoid risky behaviour in traffic situations	PART	67	2,10	0,84
	SL	158	2,14	0,55
	TI	113	**2,26**	0,51
	TR SL	45	2,20	0,55
	TR TI	5	**2,20**	0,45
	MULTI	8	2,38	0,52
	total	396	2,18	0,60
Estimate my own strengths and weaknesses	PART	67	**2,21**	0,73

[33] ($F_{(5, 390)} = 6,28$, $p < .01$)
[34] ($F_{(5, 390)} = 6,20$, $p < .01$)
[35] ($F_{(5, 390)} = 3,21$, $p < .01$)
[36] ($F_{(5, 390)} = 3,55$, $p < .01$)

What I know or can do better after FSF:		N	X	SD
	SL	158	**2,20**	0,59
	TI	113	**2,17**	0,53
	TR SL	45	**2,24**	0,61
	TR TI	5	1,80	0,45
	MULTI	8	**2,38**	0,52
	total	396	2,20	0,60
Anticipate dangerous situations	PART	67	2,15	0,72
	SL	158	2,18	0,54
	TI	113	2,16	0,56
	TR SL	45	**2,31**	0,51
	TR TI	5	**2,20**	0,45
	MULTI	8	**2,63**	0,52
	total	396	2,19	0,58
Avoid dangerous situations	PART	67	2,03	0,74
	SL	158	2,10	0,56
	TI	113	2,16	0,54
	TR SL	45	2,11	0,49
	TR TI	5	1,80	0,45
	MULTI	8	**2,25**	0,71
	total	396	2,11	0,58
Master dangerous situations [37]	PART	67	**2,19**	0,74
	SL	158	2,11	0,50
PART > TR TI ($p < .05$)	TI	113	1,96	0,67
	TR SL	45	1,96	0,60
	TR TI	5	1,20	0,45
	MULTI	8	1,50	0,53
	total	396	2,04	0,63
Estimate my own abilities in traffic	PART	67	1,96	0,75
	SL	158	2,05	0,55
	TI	113	1,96	0,52
	TR SL	45	2,13	0,50
	TR TI	5	2,00	0,00
	MULTI	8	2,13	0,35
	total	396	2,02	0,57
To have safety in mind while driving	PART	67	2,12	0,83
	SL	158	2,07	0,62
	TI	113	2,12	0,55
	TR SL	45	2,09	0,56
	TR TI	5	1,60	0,55
	MULTI	8	2,13	0,64
	total	396	2,09	0,63
Anticipatory driving	PART	67	1,87	0,81
	SL	158	2,11	0,52
	TI	113	2,06	0,56
	TR SL	45	2,11	0,53
	TR TI	5	1,80	0,45
	MULTI	8	2,25	0,46
	total	396	2,05	0,60
Braking	PART	67	**2,45**	0,76
	SL	158	**2,51**	0,61
	TI	113	**2,58**	0,51

[37] ($F_{(5, 390)} = 5,00$, $p < .01$)

What I know or can do better after FSF:		N	X	SD
	TR SL	45	**2,29**	0,66
	TR TI	5	**2,40**	0,55
	MULTI	8	**2,63**	0,52
	total	396	2,50	0,62
Driving in curves	PART	67	**2,19**	0,76
	SL	158	**2,34**	0,60
	TI	113	**2,28**	0,57
	TR SL	45	2,13	0,55
	TR TI	5	1,80	0,84
	MULTI	8	2,25	0,71
	total	396	2,27	0,63

The greatest agreement on the participants FSF outcome within the six groups turns up with "braking"; enhancement of braking skills can be found in the "top 5 improvement topics" of all six groups. Five groups agree that the participants will "estimate their strengths and weaknesses better" after the FSF training (all groups but track instructors), also five out of six groups agree that the participants will better "respond and react to unsafe behaviour of passengers in the car" after the FSF training – astonishingly all groups (all trainers and authors) except for the participants themselves reported this topic in the "top 5"; the participants themselves reported only a slight improvement in their abilities to deal with disturbing passengers. On the other hand two topics can be found in the participants' "improvement top 5" that are considered to be only of minor importance by all other groups. These are "vehicle control" and the "ability to master dangerous situations". The authors particularly reported that they thought the participants would rather recognise and avoid dangerous situations than be trained to master them. From the authors' standpoint this is logically consistent because it must be a goal of the training to avoid giving participants the impression that they are better at mastering dangerous situations after FSF. On the other hand the participants have a rather practical view on their outcomes, so they might prefer the term "master dangerous situations".

Allocation of the FSF programme topics to the four levels of the GDE matrix (concept transfer from the authors to the participants)

All six groups were asked to allocate the main topics of the FSF seminar to the labels of the four levels of the GDE matrix. The four levels were specified in the questionnaire with the following key words:

1. vehicle operation
2. control of driving situation
3. being aware of the purpose of driving and the driving conditions
4. being aware of personal attitudes and goals connected with driving

As mentioned above, the FSF programme's aim was to raise awareness of the influence of one's attitudes on safe driving mainly located on the higher GDE matrix levels, in particular level three ("being aware of the purpose of driving and the driving conditions") and four ("being aware of personal attitudes and goals connected with driving"). We used the handy four level descriptions of the GDE matrix to see if the authors' concept was transferred to the participants in the way it was intended. Herewith we wanted to verify the perception of different FSF programme parts regarding its content of teaching. For example, if the authors designed a programme part with the goal to enhance "awareness of the purpose of driving" and the participants classify the same topic under "driving control" or "vehicle operation", the concept described and the methods specified in the manual - and actually used by the trainers in FSF - might need to be revised because it did not generate the expected effects.

To compare the classification between the six groups, means are calculated and Table 48 gives an overview of the average classification in the six groups. Significant differences ($p < .05$) between the groups calculated with ANOVA are shaded grey. Details on the multiple group differences checked with Games-Howell test are shown in Table 49.

The allocation of the item numbers in the first column in Table 48 was made as follows:

1 report and discussion of past driving experiences
2 different risks linked to different driving purposes
3 changes in driving behaviour since passing the driving test
4 experiences with difficult or dangerous situations
5 effects of passengers, feelings, and distraction on driving behaviour
6 avoidance of drugs and alcohol when driving
7 intentions, behavioural strategies for future driving
8 monitoring of one's own driving behaviour and feedback from passengers
9 observing and practising difficult traffic situations
10 environmentally-sound driving and energy-saving driving guided by the instructor
11 report of previous braking experiences
12 practising emergency braking
13 emergency braking with passengers
14 observation of braking distance extension at faster speeds
15 emergency braking with surprise effects or on slippery road
16 driving in curves with and without comfortable speed
17 driving in curves with passengers

Table 48: GDE classification in the six groups (1 = level "vehicle operation", ..., 4 = level "personal attitudes and goals" (means and standard deviations) in every group

Item	PART (N = 48)		SL (N = 107)		TI (N = 102)		TR SL (N = 42)		TR TI (N = 5)		MULTI (N = 8)	
	X	SD	X	SD	X	SD	X	SD	X	SD	X	SD
1 [38]	3,10	1,17	2,93	1,14	3,41	0,95	3,05	0,94	3,60	0,55	3,50	0,53
2 [39]	2,88	0,73	2,96	0,67	3,12	0,53	3,21	0,61	3,20	0,45	3,38	0,52
3	2,96	1,25	2,85	1,21	3,13	1,21	3,24	1,12	4,00	0,00	3,75	0,71
4	2,44	0,94	2,50	0,99	2,63	1,06	2,48	0,92	3,00	1,00	2,38	0,74
5	2,81	1,00	3,07	0,94	3,02	0,91	3,21	0,84	3,40	0,55	3,38	0,74
6 [40]	3,44	0,90	3,80	0,48	3,67	0,69	3,71	0,64	3,80	0,45	4,00	0,00
7 [41]	3,21	0,90	3,38	0,92	3,67	0,67	3,45	0,77	3,60	0,89	3,75	0,46
8	2,63	1,18	2,74	1,13	2,97	1,13	2,88	1,13	3,80	0,45	2,75	1,16
9	1,79	0,74	1,96	0,88	1,92	0,86	2,05	0,76	1,80	0,45	1,75	0,46
10	2,08	1,15	2,01	1,15	1,76	1,07	1,90	1,14	2,20	1,64	1,88	1,36
11 [42]	2,10	1,12	1,63	0,98	2,37	1,24	1,60	1,04	3,00	1,00	2,13	1,13
12	1,73	0,89	1,39	0,81	1,57	0,85	1,36	0,82	1,60	1,34	1,50	1,07
13 [43]	1,96	0,90	1,75	1,04	2,06	1,01	2,14	1,07	2,40	1,14	2,88	0,83
14 [44]	2,44	1,11	2,29	1,17	2,90	1,08	2,81	1,17	3,60	0,89	3,50	0,93
15 [45]	1,90	0,78	2,10	1,05	2,34	1,09	2,33	1,16	3,40	0,89	3,13	0,83
16 [46]	2,25	1,08	2,40	1,03	2,41	1,02	2,60	1,15	3,60	0,89	3,50	0,93
17 [47]	2,27	1,07	2,41	1,05	2,71	1,10	2,81	1,11	3,60	0,89	3,63	0,74

[38] ($F_{(5, 306)} = 2,75$, $p < .05$)
[39] ($F_{(5, 306)} = 2,48$, $p < .05$)
[40] ($F_{(5, 306)} = 2,54$, $p < .05$)
[41] ($F_{(5, 306)} = 2,67$, $p < .05$)
[42] ($F_{(5, 306)} = 6,81$, $p < .01$)
[43] ($F_{(5, 306)} = 2,90$, $p < .05$)
[44] ($F_{(5, 306)} = 5,36$, $p < .01$)
[45] ($F_{(5, 306)} = 4,11$, $p < .01$)
[46] ($F_{(5, 306)} = 3,41$, $p < .01$)
[47] ($F_{(5, 306)} = 4,44$, $p < .01$)

Table 49: Multiple analysis of group differences in the mean allocation of the FSF programme topics to the GDE matrix, results of Games-Howell test

Topic	Sign. diff. between group means
1. report and discussion of the past driving experiences	SL < TI ($p < .05$)
2. different risk with different driving purposes	-
6. avoidance of drug and alcohol when driving	PART < MULTI ($p < .01$) SL < MULTI ($p < .01$) TI < MULTI ($p < .01$)
7. intentions, behaviour strategies for future driving	PART < TI ($p < .05$)
11. report of previous braking experiences	SL < TI ($p < .01$) TI > TR SL ($p < .01$)
13. emergency braking with passengers	SL < MULTI ($p < .0.05$)
14. observation of braking distance extension with faster speed	SL < TI ($p < .01$)
15. emergency braking with surprise effects or on slippery road	PART < TI ($p < .05$) PART < MULTI ($p < .05$)
16. driving curves with and without comfortable speed	PART < MULTI ($p < .05$)
17. driving curves with passengers	PART < MULTI ($p < .01$) SL < MULTI ($p < .05$)

The GDE allocation profiles of all six groups are shown in Figure 17.
Overall the mean trends seem to be quite parallel, but more important is the range shown for the respective estimations of the different groups. Outliers are shown for the trainers of the instructors' estimation for the observation driving part "monitoring one's own driving behaviour and feedback from passengers" (item No. 8), which the trainers of the instructors classified near level 4 while all other groups classified them lower than level 3. This difference in opinion was not significant because the allocation of the instructors' trainer is based on the rating of only five persons. The "report of previous braking experience" (item No. 11) was classified near level 3 by the instructors' trainers while the other groups allocated this topic around level 2. Obvious differences between the groups can be seen for the track training items concerning braking and driving curves. These topics from the track training part of the FSF programme are classified between level 3 and 4 by the instructors' trainers and the authors. The four other groups (in particular the participants and the track instructors) allocated these "practice" FSF parts somewhat lower between level 2 and level 3. This indicates that further efforts in the instruction of the track instructors need to be undertaken to convey the intended concepts of the track-training parts "braking" and "driving curves" – as defined by both the authors and the trainers of the instructors – to the track instructors (and in conclusion to the participants). It seems as if there is a considerable difference in the perception of the content of teaching for the practical training parts of FSF reported by the authors and trainers of the track instructors on one hand, and the track instructors and participants on the other.

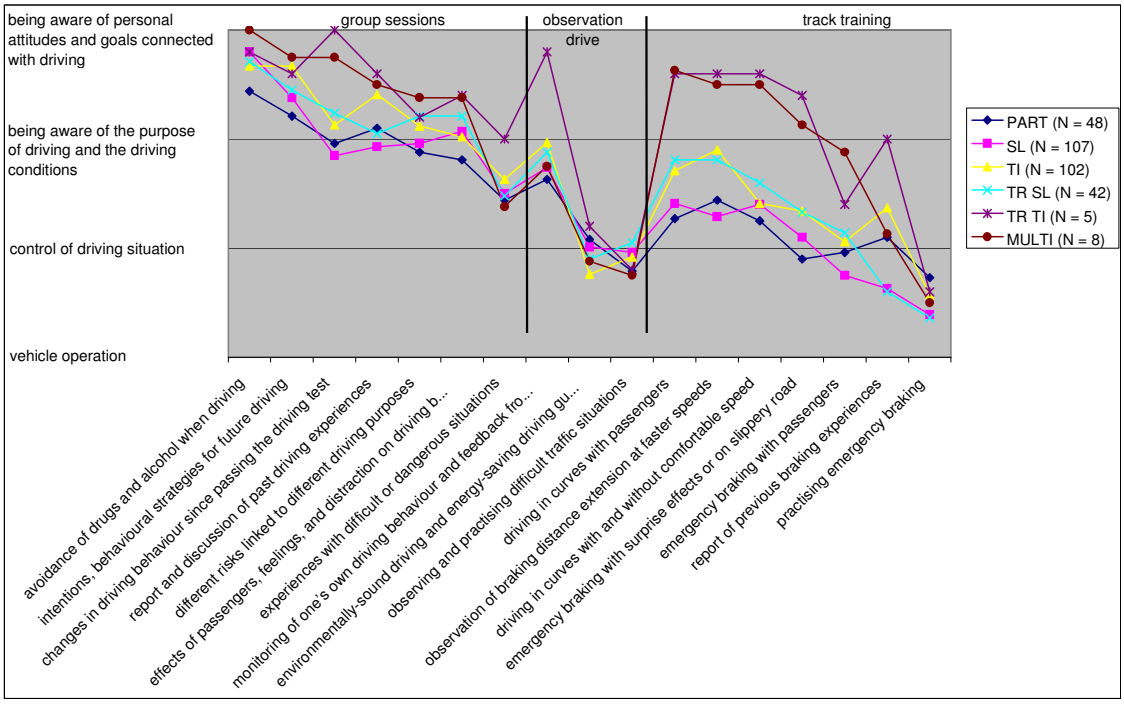


Figure 17: Profile of the allocations of FSF programme topics to the four levels of the GDE matrix; sorted by the authors' estimations within the respective topics of group sessions, observation drive, and track training

6 Discussion and conclusions

The present situation for the FSF second-phase training in Germany can be summarised as follows:
The demand for the FSF-model in the population of novice drivers is very small at this stage. DVR and other institutions trained a sufficient number of seminar trainers and track instructor to launch the FSF programme all over Germany. Driving instructors invested money to become seminar leaders and now they might be quite disappointed about the low demand for the FSF-model. However, attending the course to become a FSF seminar leader is recognised as mandatory further training for driving instructors. The training investment therefore served an important secondary purpose.
We interviewed 246 seminar leaders and only 11 of them (less then 4.5 %) had conducted FSF training so far. Many seminar leaders as well as track instructors called us by telephone after they received the questionnaire and asked if they should fill out the questionnaire despite not having implemented any FSF training yet (and, indeed, we encouraged them to fill out the questionnaire). Many of them were disappointed about the low demand for FSF, even though they did a lot of advertising in their driving schools or even motivated their former learner drivers by sending out information letters and FSF flyers. On the phone the predominant opinion of the driving instructors was that the FSF training is too expensive for young people in comparison to the benefit they get from FSF (reduction by one year of the probation period). Many of the FSF trainers commented that a car insurance bonus for FSF participants (driving beginners pay a high insurance premium) in addition to the shortened probation period could possibly encourage more participation in FSF.

Nevertheless, those novice drivers who participated in FSF are quite content with the training provided. The majority of the participants in our study judged the training to be "very good" or "good"; in general 84 % of them agreed with this positive rating. The cost-benefit-ratio was

estimated to be "adequate" by 42.1 % of the participants, but about 1/3 of the participants thought the training course was too expensive compared to the improvement they obtained.

The majority of the participants took part in FSF mainly because they wanted to shorten their probation period (this coincides with the trainers' and authors' appraisal). About ¼ of the participants reported they took part on FSF to shorten their probation period from 2 years to 1 year; another ¼ reported the same motivation but to shorten the period from 4 years to 3 years (therefore these traffic offenders are already overrepresented among the novice drivers in the FSF programme, because only about 5 % of all novice drivers in Germany have had their probation extended from 2 to 4 years). Thus about half of the participants reported "shorten probation period" as their motivation to participate in FSF. But later on the participants reported other reasons – such as a high motivation to learn more about safe driving or a desire to improve their driving skills – which can be interpreted as intrinsic motivation. In contrast, the trainers and the trainers of the trainers rather expected an extrinsic motivation for FSF participation when they indicated that the participants probably take part because participation was sponsored or the novice drivers' parents or relatives had "encouraged" it.

Looking at the implementation of FSF as reported in our survey, we can conclude that the programme is performed largely in the way the authors intended it. The activities reported by the participants and the activities expected by the programme authors and trainers coincide largely as shown in Figure 15 and Figure 16. This similarity in trends should not hide the significant differences in the assessment of activities shown in Table 45, but we think that these differences are rather a result of different reference frameworks than of actual differences between the groups interviewed. We asked: "How do you rate your own activity in the event?" with the help of a six-step scale with two fixed points, namely "does not apply at all" and "fully applies". Accordingly, differences on averages between the groups can be interpreted as different framework rather than as different judgements. The parallel trends indicate the agreement on activities shown or expected in FSF more than the accomplished mean comparisons between the different groups.
Another indicator that the FSF concept has been implemented accurately as specified by the programme authors is the consensus found in the importance given to the diverse FSF programme parts. The topics evaluated as most important in FSF by the authors are also found in the participants' as well as the trainers' high-importance-ratings. Following these results, the programme topics the authors attach importance to are considered in the same way by the trainers and participants. This should not conceal differences in the assessment of importance between the groups. In particular, attention should be given to some significant differences with the importance ratings of authors and participants: the combined importance ratings of the (three) FSF group discussions in particular differ significantly between these two groups. The authors attach much more importance to group discussions than the participants do. It is no surprise that the importance ratings of the personal strategies the participants should develop during FSF group discussions also differs significantly between authors' and participants' assessments: the development of strategies for safe driving is the most important FSF aim of all for the authors, while according to the participants it is ranked on position 13 out of 17. This indicates that in future the seminar leaders should emphasise more the importance of developing and transferring strategies for safe driving during the group discussions than they have so far.

Looking at the acquired outcomes of the FSF programme, 7.4 % of the participants reported that they have not developed any strategy for safe driving within FSF. The majority developed two or more strategies and used them in everyday driving at least once per month; the majority reported the use of strategies several times per week. However, the main aim of FSF seems to have been realised: the participants developed strategies for safer driving and apply them frequently in every day driving (even though they did not attach a great importance on the development of strategies during the training, as described above. This can be interpreted as a shortcoming of the seminar leaders who possibly could not impart the importance of the strategies for safe driving which was expected by the authors).

Regardless of this overall positive outcome it should be considered that the participants mentioned "safe vehicle control" and "master dangerous situations" in their "top 5 improvement list". The participants indicate that they got the impression they have increased their abilities to master difficult or dangerous traffic situations after the training. This – as a matter of course – cannot be the goal of any training for novice drivers, because overconfidence in one's own ability to master dangerous situations is counterproductive to road safety and this phenomenon should be monitored for coming FSF training over time.

It would be interesting to see if the novice drivers that participated in FSF are actually safer drivers and prove themselves in everyday traffic compared to comparable novice drivers without FSF experience. A follow-up study of accidents and traffic delinquency to analyse the long-term effects of FSF would be advisable. Questions about the possible safety effects of the FSF-model will be investigated in the forthcoming evaluation work which will be set up by BASt in accordance with the legal provisions of the FSF-model.

The examination of the assignment of FSF contents to the labels of the four levels of the GDE matrix reveals a certain level of adequacy. The practical parts of FSF track training (item No. 8 – 17) are mostly assigned to the lower GDE levels while the parts from the group discussions (item No. 1 – 7) are assigned to higher levels (see Figure 17).
The lowest levels overall (about level 2, i.e. "control of driving situation") are associated with the FSF parts from the track training and the observation and practice drive: all involved persons agree that "practice emergency braking" (item No. 12) and "practising difficult situations" (item No. 9) can be assigned between the levels 1 ("vehicle operation") and 2 ("control of driving situation"). The sample also widely agrees that "environmentally sound driving" (item No. 10) can be assigned to the "control of driving situation" level. For the remaining practice parts of the FSF track training the outcome is very interesting: all other track training contents that deal with braking and driving in curves (item No. 14-17) are classified between level 3 and 4 by the authors and trainers of the track instructors, while the track instructors and the participants assigned these topics to the "driving control" level. The participants do not assign these exercises to the awareness of the influence of attitudes level, rather they experienced them as vehicle control practice. As a consequence, the instruction of the track instructors for the practical FSF training needs to be reconsidered, because the authors and trainers of the track instructors pursue a different goal with the practical FSF exercises then was perceived and implemented by the track instructors and assumed by the participants.
The highest levels overall are associated with avoidance of drug and alcohol (item No. 6) and the behaviour strategies for future driving (item No. 7). This can be seen as a coherent appraisal of the seminar parts to the theoretical background of the FSF programme. The authors' intended goals of these programme parts are "attitude changing" and "reflection on driving conditions" which has been affirmed by the participants' allocation of these topics to the higher GDE levels.

Finally we can conclude that although not many courses have taken place so far, the implementation of the FSF seminar seems to be quite successful, as far as adequacy and concordance between pedagogical concept, the GDE matrix as a theoretical reference model, and practice is concerned.
We found consensus amongst the different groups with regard to the ratings of the activities as well as to the intended importance of the FSF programme parts. The importance of the main goal – the development of strategies for safe driving – should be addressed in further training courses for the seminar leaders because the seminar leaders as well as the participants assessed this topic's importance explicitly lower than the programme authors did. Further on the reported participants' improvement of their abilities to master dangerous situations should be monitored over time, for this change in participants' attitude does not assist the expected betterment of road safety.
To improve the impact of the track practice for safe driving the track instructors should be better familiarised with the goals of the FSF track training during their training. Our survey showed

that they implement the practices rather with the learning target to improve driving control than to generate awareness of the influence of attitudes towards safe driving (as intended by the programme authors and taught by the trainers of the track instructors).
More reflection should also be given on how the motivation of novice drivers to participate voluntarily in FSF could be enhanced. A good approach might be to adjust the price for FSF or to extend the benefits novice drivers get from participation, i.e. the possible reduction in insurance premium associated with FSF participation should be reconsidered.
Further efforts to evaluate the long-term effects of the training event on FSF participants compared novice drivers who have not participated should be aspired to. BASt already projects advanced studies on this subject.

7 List of literature

- Bartl, G. (Ed.) (2000). *DAN-Report – Results of EU-Project: Description and Analysis of Post Licensing Measures for Novice Drivers.* Wien: Kuratorium für Verkehrssicherheit.
- Ebel, R., Fingskes, M., Plewka, M. & Schulte, K. (2003). *Freiwilliges Fortbildungsseminar für Fahranfänger (Für sicheres Fahren). Handbuch II: Moderatoren der praktischen Fahrsicherheitsübungen.* Bonn: Deutscher Verkehrssicherheitsrat.
- Evers, C. (2000). Post Licensing Measures in Germany. In G. Bartl (ed.), *DAN-Report – Results of EU-Project: Description and Analysis of Post Licensing Measures for Novice Drivers* (pp. 89-118). Vienna: Kuratorium für Verkehrssicherheit (KfV).
- Fingskes, M. (2003). *Freiwilliges Fortbildungsseminar für Fahranfänger (Für sicheres Fahren). Handbuch I: Seminarleiter in Fahrschulen.* Bonn: Deutscher Verkehrssicherheitsrat.
- Hatakka, M., Keskinen, E., Gregersen, N.P. & Glad, A. (1999). Theories and aims of educational training measures. In S. Siegrist (ed.), *Driver Training, Testing and Licensing – towards theory-based management of young drivers' injury risk in road traffic.* Results of EU-project GADGET, Guarding Automobile Drivers through Guidance, Education and Technology, Work Package 3. Bern: Bfu report No 40.
- Jacobshagen, W. (1997). *Nachschulungskurse für alkoholauffällige Fahranfänger (NAFA) – Kurspraxis, Wirksamkeit und Akzeptanz.* Köln: Verlag TÜV Rheinland.
- Jacobshagen, W. (1998). *Nachschulungskurs für alkoholauffällige Fahranfänger nach dem Modell NAFA in Deutschland: Klientel, Kursdurchführung, Wirksamkeit und Akzeptanz.* In: Driver Improvement 6. Internationaler Workshop. Bergisch Gladbach: Berichte der Bundesanstalt für Straßenwesen (BASt), Heft M93.
- Kiegeland, P. (1999). *Wirksamkeitsuntersuchung zum Sicherheitstraining für PKW-Fahrer und –Fahrerinnen.* Schlußbericht zum BASt-Forschungsprojekt FE82.051/1993, unpublished.
- Kirk, R.E. (1995). *Experimental Design: Procedures for the Bahavioral Sciences* (3rd ed.). Pacific Grove: Brook/Cole.
- Schulz, S.-O., Henning, H.J. & Chaselon, F. (1995). *Jugend fährt sicher.* Schlußbericht zum BASt-Forschungsprojekt 2.9124, unpublished.
- Seidel, E. & Beetz, U. (1978). *Wirksamkeitsanalyse zum Sicherheitstraining des Deutschen Verkehrssicherheitsrates.* Bergisch Gladbach: Schriftenreihe der Bundesanstalt für Straßenwesen (BASt), Unfall- und Sicherheitsforschung Straßenverkehr, Heft 18.
- Siegrist, S. (ed.) (1999). *Driver Training, Testing and Licensing – towards theory-based management of young drivers' injury risk in road traffic.* Results of EU-project GADGET, Guarding Automobile Drivers through Guidance, Education and Technology, Work Package 3. Bern: Bfu report No 40.

- NETHERLANDS

NovEV
Results of the Dutch second phase

Saskia de Craen, Jan A.M.M. Vissers, Maura Houtenbos and Divera Twisk

Executive summary

Participants
After an appeal by mail and telephone, 376 young novice drivers agreed to participate in the project. Unfortunately, during the course of the project, many of the participants dropped out. Out of 376 young drivers that initially agreed to participate, only 127 (33%) completed all parts of the project.

The participants who did not want to participate, those who dropped out, and those who finished all parts of the project were compared for a number of variables. This led to the conclusion that there was no major problem with selective drop-out. Naturally, the groups did differ on at least one aspect, namely for one reason or another some completed the project and others did not.

Training programme + objectives
The second phase training consisted of the following modules:

- *An on-road feedback drive*
  The objective of the feedback drive was to present the driver with feedback about his driving performance. It was different from instruction drives, as the instructor confronted the driver with his "expert" observations in order to make the participant "think" and reflect. So he did not tell the participant what to do, but encouraged him to draw his own conclusions. During the first feedback drive the participant and instructor were accompanied by a second participant who rode along as a passenger. The drive was followed by a discussion between instructor, passenger and driver.

- *Training on a closed track*
  The objective of the track training was for participants to experience the limits of their skills in vehicle control and to share these experiences with other group members.

- *A group discussion*
  The objective of the group discussion was to stimulate recognition of potentially hazardous situations in rather "normal" social situations. The discussion was based on video sketches, depicting typical situations (incidents rather than accidents) involving young drivers (men and women). The moderator encouraged the youngsters to reflect on the events.

- *An evaluation on-road feedback drive (about a month later)*
  The objective of this second feedback drive was the same as the first feedback drive, that is to present the driver with feedback about his driving performance.

Evaluation design and data collection methods
The effect of the track training and group discussion was studied using a before-and-after design with a control group. Participants were randomly assigned to the control or the experimental (treatment) group. The *control group* participated in both feedback drives. In addition to the feedback drives, the *experimental group* also participated in track training and in a group discussion.

Table 1 Evaluation design and data collection methods

	Training programme		Instruments
	Experimental	*Control*	
December 2003 **Pre-test** *One month before training*	Questionnaire	Questionnaire	**Questionnaire** Contained items on risk awareness, self-assessment of skill, and situation judgements
January 2004 **Training day**	Pre-test feedback drive	Pre-test feedback drive	**On-road observation form** An assessment tool to describe the driving performance of a driver. The driver himself and the driving instructor completed these forms after the feedback drive. **Driving Assessment** Assessment by the instructor of the quality of driving in three fields: vehicle control, driving skills, and calibration skills
	Track Exercises		
	Group discussion		
February 2004 **Post-test** *One month after training*	Questionnaire	Questionnaire	**Questionnaire** Contained items on risk awareness, self-assessment of skill, and situation judgements
	Post-test feedback drive	Post-test feedback drive	**On-road observation form** An assessment tool to describe the driving performance of a driver. The driver himself and the driving instructor completed these forms after the feedback drive. **Driving Assessment** Assessment by the instructor of the quality of driving in three fields: vehicle control, driving skills, and calibration skills **Satisfaction questionnaire** This questionnaire contained questions on how satisfied participants were about the different components of the training day and the feedback drives.

Results & Conclusions by instrument

Satisfaction questionnaire
Young drivers were not motivated to participate on a voluntary basis in a second phase training. However, once in the course, novice drivers were enthusiastic about the training day. Within the training day, the group discussion was rated as the least attractive part, while the feedback drive was about as attractive and useful as the track training. The message of the second-phase training was well-understood. There were no indications that the young, novice drivers overestimated their skills, as a result of the training.

Questionnaire
The questionnaire contained items on risk awareness, self-assessment of skill and judgements of traffic situations on photo. The results from the questionnaire are somewhat unclear; some effects of the training were found, but not consistent and not always in the expected direction.

In line with expectations, the items concerning risk awareness confirmed that young drivers do not seem particularly concerned in general, and especially not about driving too fast. A least 60% of the respondents are not concerned about driving too fast. On the other hand, it turned

out that young drivers are, overall, rather confident about their driving skills. At least 30% of the participants believe they are (very) strong in all skills, and in some skills more than 60% believe they are (very) strong.

It was expected that these opinions would improve as a result of the training day. Detailed analyses showed no effect of training on these variables. Further research is needed to demonstrate that the questionnaire itself is sensitive enough to register changes as a result of a short term intervention. The fact that there were significant gender differences in these issues, led to the conclusion that this part of the questionnaire possibly measures more stable attitudes or personality traits (which could not be changed with a one-day training course or within the period of a month).

On-road observation form
After the feedback drive, an on-road observation form was filled out by both the instructor and the participant, which contained items on driving skill and assessment of complexity of the driving task. The young drivers' assessment of their own driving skills and task complexity did not change as a result of training. This implies, that the objective of the course to inform young drivers about their limited skills and the high complexity of the traffic situation did not result in a more cautious self-estimation. On the positive side, this result indicates that the training day and more in particular the track training did not lead to a higher estimation of skills and a lower estimation of the complexity of the driving task.

To study the accuracy of the driver's self image, their self-estimation scores were compared with the instructor's assessment of the young driver's competencies. On "vehicle control and general skills", instructors and participants did not differ in their assessment neither on the pre-test nor on the post-test. As expected on "safe and defensive driving" in the pre-test, participants rated their performance higher than the instructor did. As the course was directed at improving self-assessment skills, it was expected accuracy to improve in the sense that their assessment would be more in line with that of the instructor after the training. This was not the case.

Generally, from the results from the on-road observation form, it can be concluded that while the instructors did see some improvement as a result of the training, the participants did not.

Driving Assessment
Task conditions between control group and experimental group differed systematically on the pre-test. Therefore, it cannot be excluded that the observed difference in task performance between control group and experimental group is a reflection of these test conditions rather than a significant difference between the two groups.

Within the experimental group, the performance of the participants of the two different training locations differed significantly. This, despite the fact that at both locations the participants had received exactly the same training (on paper). Where performance at location A was improved by training, driving performance at location B got even worse. Because the test conditions for the participants of the two locations were the same, this result is reliable.

The process evaluation indicated that despite their organisation's involvement in the NovEV project, the trainers from location B did not share the same opinion on the definitions of a "useful" training. As a result, these trainers had to give a type of training they did not believe in. This could have (subconsciously) affected the way they gave the training, or the way the participants perceived the training. Research has shown (ADVANCED, 2002) that any education, looses its strength if the educator is not absolutely convinced about what he/she is teaching. Moreover, that the effectiveness of the education is largely dependent on the person, the beliefs of the teacher, and his behaviour (Hale and Glendon, 1987). For a more detailed discussion of the role of the "teacher", see the ADVANCED report.

General conclusions

In the Dutch pilot, the recommendations of the ADVANCED report were closely followed with respect to the content of the course and the evaluation of its effects. However, as stated earlier, in practice these recommendations were not always followed in one of the two locations.

In this study, it has been demonstrated that, on the one hand, the second phase is recognized by the participants as a useful and necessary part of their driving career. On the other hand, the high refusal rate demonstrates that youngsters are not interested in participating on a voluntary basis. The effects of the course are limited, and can even be negative, if trainers are not fully equipped to give the course, indicating that a much greater effort is needed in training second phase trainers than has been the case in this project.

1. The projects in detail (training and evaluation)
In this section of the report, the organization, content, and results of the Dutch Pilot in the European NovEV project are presented. The Dutch report is based on two evaluations: a process and an effect evaluation. The process evaluation is reported in paragraphs 7.3 to 7.5. It deals primarily with the implementation and organization of the course. The results are based on interviews with all partners involved: organizations, trainers, instructors and researchers. This process evaluation was carried our by the Traffic Test company in the Netherlands.

The effect evaluation is reported in paragraph 7.2 with regard to the selection of the course participants and the self-selection effect, in 7.5.1. on the attractiveness of the course for the participants, and 7.6 to 8. on the effects of the course on self-assessment and driver behaviour. The effect evaluation was carried out by SWOV (Institute for Road Safety Research, the Netherlands).

1.1 Introduction
The high accident risk of young/ novice drivers has led to initiatives within Europe to find new methods for accident prevention. One such possibility is a new approach to driver training, in particular to post-license training. In the European project ADVANCED the basic principles of a successful advanced driver training are described. The ADVANCED project concluded that the primary objective of advanced driver training is to enhance and stimulate the development of higher order skills. These skills are related to hazard perception, self-assessment and situation awareness, and are known to develop relatively slowly in comparison to other driving related skills such as vehicle handling and the mastery of traffic situations.

Many studies indicate that there is a relation between hazard perception and self-assessment of skills. For example Brown (1989) claims that the perception of risk cannot be studied in isolation of both these elements. This balance between hazard perception and self-assessment has been called calibration. Calibration is seen to be an essential element in safe driving. At any moment in time, a driver needs to be actively engaged in assessing what the driving task requires in terms of actions or the avoidance of actions, and the potential difficulties involved (Kuiken & Twisk, 2001).

In brief, there are indications that young drivers underestimate the risk of an accident in a variety of hazardous situations At the same time there seems to be a problem with the assessment or evaluation of one's one driving skills. For example, the young driver underestimates what is needed to cope with a dangerous situation, they overestimate their own driving skill (Deery, 1999). McKenna et al. (1991) concluded that this overestimation of driving skills is caused by a "positive self" rather than a "negative other" bias. This could be caused by the fact that the young driver has encountered only a limited number of critical traffic situations, which may provide a false sense of mastery and safety. Therefore, a 2nd phase driver training could be used to eliminate or reduce this false sense of mastery and safety. In other words, improve calibration skills.

There is evidence to suggest that especially the sub-standard levels of higher order skills in novice drivers are one of the main causes of their increased crash rate (Deery, 1999; Willems & Cuyvers, 2004; Engström et al., 2003). In the basic (pre-licence) training phase, drivers are trained with respect to vehicle control and the mastery of traffic situations. What is essential in this phase is the faultless and automatic application of such driving routines. Of course, issues such as hazard perception and risk awareness are addressed, but as driving experience is still very limited at this stage, the effect is probably relatively small. After completing his/her basic driving course, and passing the exam, the novice driver gains experience, but also is exposed to new risks. After 6 months of independent driving, he has driven in “unfamiliar” situations, encountered new traffic situations, has started to develop his own driving style and to regard car driving as a means to an end (e.g. to go to a party to have fun) rather than as a meaningful activity in itself.

These new developments in the novice driver's career calls for a second phase in driver training. The objective of the second phase is to address these experiences and to contribute to the prevention of the associated risks. Furthermore, it is essential that any training should avoid overconfidence to develop. Research findings suggest that advanced multi- phased training for novice drivers that focuses on vehicle skills like skid control and emergency manoeuvring skills is counterproductive. Therefore, any second phase training needs to ensure, that such overconfidence does not result from training. This creates a dilemma for the second phase training. On the one hand, in order to be effective it is important for participants to be highly motivated and to find the courses attractive and stimulating. On the other hand the type of training that is most attractive (namely vehicle handling skills like skidding) should be excluded from the course.

To conclude, the effectiveness of second phase of driver training depends on:
- The adequacy of the training module to stimulate and to enhance higher order skills,
- The timing of the second phase in the total learning process,
- The extent to which the chance of overconfidence developing is minimized,
- Its attractiveness for the target group.

The question that also needs to be answered is the relationship between the quality of the basic driving course and effectiveness of the second phase. Insight in this relationship is currently missing. On the one hand, it can be reasoned that a poorly educated driver does not benefit from advanced driving courses because of his poor driving routines. On the other hand, it seems likely that poorly educated drivers benefit the most, because for them there is still a lot to learn.

1.2 The content of the NovEV Dutch second phase

In the Netherlands, driver training consists only of a basic driving phase. The content and training methods are not standardized, and it is left to the driving instructor to decide on how and what to teach. Driving standards are ensured by the content, reliability, and validity of the compulsory driving test. The government sets these standards.

As discussed above, the impact of the basic driving phase is limited and for this reason the Netherlands is participating in the NovEV project in which experts in the field have designed, implemented and evaluated a second-phase course.

This course had the following objective: *"To enhance self-assessment skills, risk and safety awareness by feedback and training/coaching with respect to an individual's driving behaviour, personal style and decision-making characteristics"*.

The structure and content of the training course followed closely the best practice recommendations of ADVANCED (pp. 134-138), and consisted of the following elements:
- An on-road assessment drive (first feedback drive)
- Training on a track
- A group discussion
- An on-road evaluation drive (second feedback drive; about one month later)

This training took place at two locations:
- The ANWB (Dutch Automobile Club) track training site situated near Lelystad: in this report it is frequently referred to as "Lelystad"
- VVCR (post-licence driver training centre) track training site situated near Rijssen, referred to in this report as "Rijssen"

In both locations, the course structure and content was the same, and a detailed outline was described in the blueprint. However, differences did occur in the actual implementation and execution of the course. More detail on this can be found later in this report.

The exact content of the training elements depended on the "specific needs" of a driver.
To assess the specific needs of a driver, instruments were developed that were used to identify their particular weaknesses. These instruments were used for diagnostic purposes by the trainer and to provide feedback to the learner driver.

1.3 Organization

The design, implementation and evaluation of the second phase course was carried out by a consortium of partners who all contributed towards the financing, organization and expertise in the project. These partners came from many fields, like exam and training centres, driving schools, and research and governmental organizations. This cooperation led to the development of a blueprint for the course, its implementation and evaluation that all partners agreed upon (see 7.3.1. for a detailed description of the organization). Choices with respect to content, organization and implementation were not only guided by the guidelines from the ADVANCED project, but also by consideration of what was feasible in a full scale implementation. For instance, although it led to a (too?) tight time schedule, 3 groups participated in each training day. In this manner, facilities such as the track were used in the most economically efficient way.
Appendix 1 contains a complete list of the participating organizations. The general coordination was carried out by the ROVG (regional road safety council of Gelderland).

1.4 Evaluation

To assess the effectiveness of the second phase training, the participants were divided into two groups. The first group (experimental group) would follow every part of the training. In other words, they would attend both feedback drives, the track training and the group discussion. The second group (control group) would only attend the feedback drives. The difference between both groups would then reveal the effect of the track training and group discussion.

In order to measure the differences between the experimental and control groups, several instruments were developed:

- Questionnaires: concerning driving skills, self-assessment and risk awareness. The participants completed the questionnaires twice, one month before and one month after the training.
- Diaries: Semi-structured questionnaires in which driving events were reported by the participants. The diaries were also completed by the participants one month before and after the training. The results of these diaries, however, exceed the scope of this report, and will not be presented here.
- On-road observation form: an assessment tool to describe the driving performance of a driver. The participant himself and the driving instructor completed these forms after the first and second feedback drive. This instrument was not only used for research purposes. By comparing the assessments, the forms were also used as input for the discussion after the feedback drive.
- Driving assessment: assessment by the instructor, based on the feedback drives, of the quality of driving in three fields: vehicle control, driving skills, and calibration skills[48].
- Satisfaction questionnaire: this questionnaire contains questions on how satisfied participants were about the different components of the training day and the feedback drives. The questionnaire was completed after the last part of the project, namely the second feedback drive.

[48] Calibration is defined as the balance between self-assessment of skill and risk awareness. A central element in this balance is the skills a driver actually has, versus the skills the driver thinks he has.

The evaluation design and measurement instruments conform to the quality criteria for evaluation research. (ADVANCED, 2002; pg 139-150).

2. Young Drivers Experience: the Dutch project

2.1 Participants / subjects

The aim was to have about 300 young novice drivers to participate in the project. Addresses of newly licensed drivers, between the age of 18 and 25, were obtained from the Central Licensing Bureau. Those that received their license in the summer of 2003 received a brochure and an invitation by mail to participate in a "challenging" (and free) safe driving course. The participant's travel expenses were refunded and, to stimulate participation, participants could win a holiday for two or free car insurance for one year by entering a lottery. As too few participants accepted the written invitation (about 10%), a new group of novice drivers were invited by telephone.

The young drivers who did not want to participate in the project (approximately 140) were asked to answer a few questions so they could be compared with the people who did want to participate, thereby checking for a selection bias. One of the questions asked concerned the reason why they did not want to participate in the project. Figure 18 shows the percentages for each reason. (The percentages do not add up to a hundred, because it was possible to give more than one reason for non-participation).

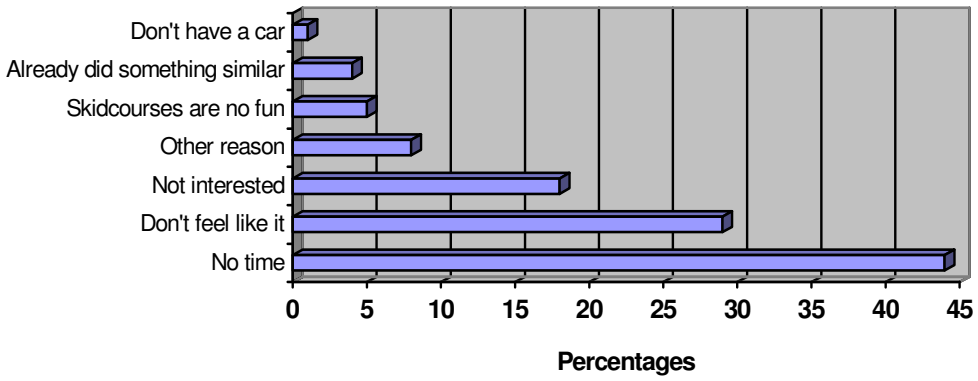


Figure 18 Reasons for not participating in the project

The most common reasons for not participating in the project were lack of time and lack of interest.

After the round of phone calls, 376 young novice drivers agreed to participate in the project. Unfortunately, during the course of the project, many of the participants dropped out. The participants were expected to come to a location twice, the first time for a feedback drive and for some participants a training (experimental group), the second time only to attend a feedback drive. Table 2 shows the attendance on both feedback drives.

Table 2 Attendance in the feedback drives

		Post training feedback drive		Total
		Absent	Present	
Pre training feedback drive	Absent	198	15	213
Pre training feedback drive	Present	36	127	163
Total		234	142	376

Of the 376 young drivers that initially agreed to participate in this project, only 127 completed all parts of the project. Most of the participants dropped out before the first feedback drive, the first time they had to come to a location. One explanation is that participants became aware of what the contents of the project exactly were. We suspect that when the participants signed up for the project during the round of phone calls, most of them thought it was a one-day skid course. When they found out that they were expected to fill in questionnaires and had to come to a location twice, they dropped out. The extreme weather conditions – there was a traffic warning issued not to go outside unless you really had to, on some training days – and the distance to the training locations (Lelystad and Rijssen –a more than one hour drive for most of the participants) could also explain the high number of absentees.

Overall, this means that out of the 500 participants that were contacted to participate, about 340 were not interested enough to participate. About 140 refused immediately, and about 200 changed their mind later on. This indicates that a 2nd phase training is not something that young, novice drivers would attend voluntarily. Of course in this project they had to do much more than just a one day training; they had to fill in a number of forms, and come to a location twice. Moreover, there was a strict timeframe when they had to attend the training; it was not even possible to attend one week later.

Selection bias

Because of the high percentage of dropouts, before and after the start of the project, it is important to realize that such a dropout can be selective, thereby causing a selection bias. If, for example, relatively more woman than men drop out or relatively more experienced drivers, the sample would not be representative anymore. To study the selectiveness, all novice drivers (whether they were willing to participate or not) were asked questions about their age, training and driving experience.

In order to estimate the selectiveness of the dropouts, five groups were compared (see Table 3). The first group in this table consists of young drivers who did not want to participate in the project when they were asked by phone (n=138). The second group agreed to participate, but dropped out before the first day, or did not show up on the first day. The third group consists of those who were present the first day, but dropped out before the second day (n=36). And the fourth group consists of those who could not make it on the first day (the pre-training feedback drive) but were motivated to come the second day (n=15). None of these groups were used in the main analyses, but their scores were used illustratively in most of the Figures. For example the group who did the first feedback drive, but was absent from the second, was used for their scores on the first feedback drive. The last group in Table 3 consists of the respondents who finished all parts of the project (n=127). This group of 127 subjects was used in the main analyses.

Table 3 Selective drop-out -- Group comparison

		Not in the study				In the study
		Did not agree to participate after phone call (N=138)	Absent – Absent (N=198)	Present – Absent (N=36)	Absent – Present (N=15)	Present – Present (N=127)
Gender	% Male	50%	48%	56%	40%	61%
Age	Mean	21	20	20	21	20
	Std.	2	2	1	2	1
Number of months drivers' licence	Mean	9	9	8	7	8
	Std.	2	2	2	2	2
Hours of training for drivers' licence	Mean	42	42	39	46	39
	Std.	18	16	17	27	17

As can be seen in the table, there is not much difference between the group of respondents that did not want to participate after the phone call, the respondents that did not finish all parts of the project and respondents who did participate in all parts. The only noticeable difference is that in the final sample of young drivers 60% is male. The results led to the conclusion that there is no problem with self-selection and selective drop-out.

It should be remembered that the respondents in the study and those who dropped out somewhere along the way, could only be compared for a couple of variables. These variables indicate that there is not much difference between the groups. Naturally, these groups do differ in at least one aspect, namely for one reason or another some completed the project and others did not.

2.2 Training programme

This section covers the (organization of) different parts of the training programme. First something will be said about the project organization in general. After that the feedback drives are discussed, followed by the track training and group discussion.

Project organization

At the onset of the project, Traffic Test drew up a blueprint for the implementation of the second phase training. All members of the consortium were involved in its development (see Appendix 1). The blueprint was continuously revised and adapted during the preparation phase.

Due to financial constraints, such as the research budget and the participation cost in the training, certain choices were made.

Important considerations include:

- With the group discussion in mind, it was initially intended (as proposed in the ADVANCED guidelines) to work with a course instructor specialised in group dynamics. This instructor would accompany the participants during the entire day and also be responsible for both the course introduction and the final evaluation. However, due to the considerable costs involved, it was decided to let the trainers of the training facilities carry out the group discussions.
- It was too expensive to have – apart from the training staff- an extra person present at the training facilities each day to coordinate all activities and solve ongoing problems.
- For economic reasons, especially concerning the use of the track, it was important to optimise efficiency in using the training facilities. Consequently, it was necessary to train three groups of participants simultaneously on one day, which resulted in a very tight schedule with little room for manoeuvre if problems did arise.

Many parties involved had a specific task to perform. Responsibility for the execution of these tasks was largely delegated to the parties involved. Consequently, each party started out with their specific task, without supervision from the project management concerning the way the tasks were executed. It is thus possible that, due to the absence of supervision from the management team, some tasks were not carried out as efficiently as others.
This structure also had a positive effects, however. All partners had to adapt the ADVANCED project recommendations to their own contributions. In this manner, the outcomes of ADVANCED had a bigger impact than it might have had with a more centralized approach involving only a handful of partners.

Pre-test and Post-test Feedback drive
The objective of the feedback drive was to present the driver with feedback about his driving performance. It consisted of a drive on public roads of different categories. It was different from instruction drives, as the instructor confronted the driver with his "expert" observations in order to make the participant "think" and reflect. So he did not tell the participant what to do, but encouraged him to draw his own conclusions.
In the pre-test feedback drive, the participant and instructor were accompanied by a second participant who rode along as a passenger. The drive was followed by a discussion between instructor, passenger and driver. As input for the discussion, the "on-road observation form" for the participant-driver was completed by the participant-driver himself, by the passenger, and by the instructor. In the post-test feedback drive, no second participant was present.

The locations of the feedback drives were rather different for the experimental group than for the control group. The feedback drive for the experimental group had to take place in the vicinity of the training location, which is located more than 1 hour's drive from their home town. So most probably, they did not drive in a familiar environment. The control group, however, performed their feedback drives in an area close to their own town. This was frequently the area in which the participant had taken driving lessons and/or their exam. During the second feedback drive, all participants (both experimental and control group) drove in an exam-area close to their town. In contrast to the initial feedback drive, individual participants were alone with the trainer. The process evaluation interviews demonstrated that, in the instructors' opinions, their assessment of the participant's driving performance was partly dependent on the area in which the feedback drives took place. In their view, it is likely that the feedback drives in familiar conditions were more positively assessed than drives that took place in more unfamiliar conditions.

Track Training
The track training consisted of the following exercises:

- ABS and non-ABS braking exercises: 30 and 50 km/h
  - Goals: Understanding the differences between ABS and non-ABS, experiencing the sensation of ABS braking, understanding the effect of speed on braking distances.
- Demonstration at 50 and 60 km/h and showing the effect on braking distance.
- Driving on to the verge
  - Goal: To experience the sensation of going on to the verge and semi-loss of control
- Aquaplaning
  - Goal: To show inability to steer when aquaplaning. Participants were inside the car when the instructor demonstrated the exercise.
- Driving around bends
  - Goal: To show how small increases in speed can cause the vehicle to slide when driving around bends
- Parallel braking exercises
  - Goal: To show how easy it is to cause a pile-up unless proper safety margins are respected.

Group discussion
The group discussion was based on video sketches. The video depicted typical situations (incidents rather than accidents) involving young drivers (men and women). It highlighted issues such as distractions: music, mobile phones, peer pressure, passers-by; multi-tasking, pressure from other drivers, tailgating (insufficient safety margins), vehicle loading, etc. The moderator encouraged the youngsters to reflect on the events. The objective was to stimulate recognition of potentially hazardous situations in rather "normal" social situations. The intention was to use the video sketches as a basis for further, more spontaneous discussion amongst the participants, led by the trainer.

2.3 Selection and training of trainers and instructors

The trainers (those who gave the track training) were employees at the track-site and experienced trainers in voluntary, post-licence driving courses.
The instructors (those who did the feedback drives) were examiners from the Dutch driver testing centre and driver instructors with extra qualifications in the field of driver training and coaching.

The instructors and the trainers were instructed separately. The following sections describe these instruction sessions.

Instructions about the feedback drives
An instruction meeting was organized for the instructors involved. During the meeting, three main items were discussed. First of all, the on-road observation forms were explained. Furthermore, instructions were provided on how to give feedback to the participants during and after the feedback drive. Finally, how to fill in the driving assessment form was explained.

Track training and group discussion
Track trainers received a one-day training on:

- The objectives and design of the project
- The execution of the track training
- The execution of the group discussion

The objectives and design of the project
The project manager gave a short explanation of the objectives and design of the project. The basic principles on which the second phase training is based were briefly discussed. Initially, the trainers gave the impression that they had received enough information to carry out the training, according to the second phase training guidelines.

However, during the actual implementation of the course, it became apparent that the trainers of one of the locations (Lelystad) were not convinced of the value of the basic principles of the course. This may have influenced the impact of the track training itself.

Instructions about the track training
A trainer of the VVCR presented the scenario for the track training at the instruction meeting. The approach of each exercise was discussed. Subsequently, a tryout was performed with the track training. Based on this tryout, several aspects of the track training were adjusted.

According to the trainers, it would have been useful to have a rehearsal of the entire second phase training. Instead, problems that occurred now had to be solved on site during the training itself. On the first training day at the VVCR (Rijssen) three members of the working party were present (project manager, representative of the SWOV, representative of Traffic Test). These persons were able to step in where necessary. Furthermore, they evaluated the course with the trainers at the end of the day. Thus, the first training day in Rijssen could be seen as a field-test. No members of the working party were present on the first day of training at Lelystad

Instructions about the group discussion
A tape with the video clips was shown at instruction meeting and it was discussed how to conduct a group discussion using these clips. During this discussion, the role of the trainer as a leader of the discussion, was entered into at great length. An important principle that was stressed was that the clips should be used to get the discussion started, and that it was certainly not the intention to show and discuss all video clips on the tape. If the discussion was running well, showing only the first clip could be sufficient.

In addition, the question of how to give feedback as a discussion leader was discussed in detail. Furthermore, the need to write general conclusions and points of view on a whiteboard was mentioned as important.
The tryout that was performed with the group discussion went well and did not lead to any adjustments of the procedure.

2.4 The process evaluation: implementing the training

The process evaluation was carried out to be able to document how successfully the blueprint was implemented and what lessons could be learned from the experiences of relevant actors. On behalf of the process evaluation three discussion meetings were arranged:

- A discussion with the driving instructors and the examiners who were involved in the feedback drives (feedback drive instructors).
- A discussion with the track trainers.
- A discussion with the researchers.

In addition to these three meetings, any other relevant information on the process was also collected. This primarily concerns subjects discussed in the meetings of the 'second phase driver training' workgroup.

The following three sections (0, 0 and 0) contain the results of these discussions. The last section (0) contains the results of a questionnaire which the participants filled in with their opinion of the training.

2.4.1 Results of the discussion with the feedback drive instructors

General observations concerning the training programme
The feedback drive instructors were very enthusiastic about the Dutch second phase training initiative. According to their experience, immediately after the driving exam, errors creep into the driving behaviour of young novice drivers. In the eyes of the feedback drive instructors, second phase training can be an effective way of correcting these errors in their driving style. They had the impression that the participants were open to critical remarks on their driving style and that they were willing to improve their driving behaviour. It is for this reason that the instructors have doubts about the representativeness of the group that took part in the experiment. In their view, the group of young novice drivers that really are a problem for road safety probably didn't participate. In their opinion, this means that the second phase training programme will have to be mandatory in order to be effective.

Strong and weak points of the feedback drives
Strong points:

- Most participants showed their normal driving behaviour. The feedback drive is not seen as a driving test and participants are not afraid to make mistakes.
- Participants were open to critical remarks and the instructors have the impression that participants are willing to make use of the advice they receive.
- The participants were very sensitive to the remarks of fellow participants. Their comments have, in general, more impact than those of the feedback drive instructor.
- Working with driver profiles (based on a questionnaire) was a good basis for discussing the strong and weak points in one's driving style.
- The overall driving performance of participants was good –to- reasonable.
- Women achieved better results in the feedback drives than men. This is different from the situation at the driving test. This may be due to the fact that the feedback drives measure different aspects of driving style than in the normal driving test.

Points that can be improved:

- Both participants and feedback drive instructors were not always fully informed about the goals and content of the project. Thus, participants that had been allocated to the control group, didn't know they would not get a road safety training on the track. This resulted in dissatisfied responses from a lot of participants. However, at the end of the feedback drive, these participants had positive opinions about the usefulness and attractiveness of the feedback drive.
- The time for discussion after the feedback drive is too short to be able to talk through all experiences. This is partly due to the fact that so much paperwork has to be done (filling out profile scores and on-road observation forms).

Circumstances during the feedback drives
When performing the feedback drives during the pre-test, the instructors were confronted with severe winter weather conditions: intensive snowfall and snow-covered roads. For many participants this was their first experience with such extreme weather conditions. According to the feedback drive instructors, this led the participants to drive extremely carefully by driving very slowly and by keeping larger safety margins than usual. According to the instructors, this influenced the way the driving behaviour of the participants was assessed. Because weather conditions were quite normal during the feedback drives during the post-test, it is more likely that participants displayed their normal driving behaviour. The differences in weather conditions between pre-test en post-test situation imply that it is difficult to compare the results of the feedback drives between pre- and post test and between experimental and control group.

Planning and organization of the feedback drives
Feedback drive instructors are rather critical about the way the feedback drives were planned. Instructors found that they were not always informed in time about the timetable and modifications in the schedule were not transmitted in time. Instructors found it very frustrating when participants didn't show up. One of the instructors went so far as phoning participants the day before they had their feedback drives to check if they were informed about their appointments and if they could keep them. This worked very well and all these participants eventually appeared.

2.4.2 Results of the discussion with the track trainers

General observations concerning the training programme
The trainers of the training sites in Lelystad (ANWB) and Rijssen (VVCR) were also very enthusiastic about the initiative to set up and evaluate the experiment with the second phase training programme. However, during the discussion it appears that in the case of the ANWB trainers there was some uncertainty about the basis assumptions of the second phase training programme, especially when it comes to the contents and the working method of the track training.

For the ANWB trainers, working with this young age group is relatively new. The trainers of the VVCR already have a long tradition in working with young novice drivers. The VVCR was also involved in the EU-project ADVANCED, and, in the framework of the 'Young Drivers Project', the VVCR has already been carrying out training programmes for young novice drivers for some years. For this reason they already are more familiar with the basic principles of the second phase training programme.

During the meeting with the track trainers the discussion focussed partly on the benefits of skill-oriented track training. What emerged was that particularly the ANWB trainers seem to have a rock-solid faith in the usefulness of their skill-oriented training. Especially when training conditions are difficult (which was the case in the Dutch experiment due to the wintry

conditions) and trainers do not have much or any experience in performing the training programme and in working with novice drivers (which applies to the ANWB trainers), trainers tend to fall back on their normal working methods. This perhaps also explains why ANWB trainers have difficulty in accepting the new principles of the second phase training programme, although initially they supported these assumptions and agreed upon the structure of the training programme. Considering the initial enthusiasm for the second phase programme, and the fact that the Advanced project philosophy of track training was never questioned during the working group meetings, the project management would only have been to predict this situation occurring in practice by conducting a full rehearsal.

What we can learn from the Dutch experience is:

- An introduction to the training programme of three hours is not enough to teach the trainers the skills and motivation necessary for an effective execution of the course, especially if trainers do not have experience in working with the target group of young novice drivers.
- It is necessary to have a more profound discussion about the assumptions of the training programme. Having trainers state that they support the training programme is not enough.
- Trainers with little or no experience in working with young novice drivers need time to build up experience with the training programme in practice.

Track training

The track trainers in general agreed upon the structure of the track training and upon the exercises that are part of the track training. In the case of the ANWB trainers, this seems to contradict the fact that they questioned the basic assumptions of the second phase training programme. Perhaps this has to do with the fact that, in essence, they support the exercises (because they are also part of their own skill-oriented training programme), but they do not yet understand the different methods involved through which exercises support the principles of the second phase.

A general comment from the trainers was that, in some cases, they would have liked to have more time available, so that participants could experience the effects of the manoeuvres more frequently, e.g. the influence of speed on the braking distance.

Due to the poor weather conditions most of the training sessions in Lelystad could not be performed according to the blueprint. Because of the exposure of participants to snow and rain, it was not possible to have a discussion with the whole group after each exercise on the track. In Rijssen, participants could stand in a bad weather shelter, so discussions were carried out according to plan.

In the case of the ANWB trainers, there was some misunderstanding concerning the degree to which they could give *instruction*. This led to the conclusion that in Lelystad there had been insufficient communication about the objectives of the exercises.

Another point of concern is that participants need to have the possibility to experience the exercises outside the car, when standing on the side of the track. In the braking exercises, for instance, the impact of speed is sometimes felt more outside than inside the car. In Rijssen the training group was always split in two: one group driving and one group observing. Due to the bad weather conditions in Lelystad, this was not the case.

Group discussion

The group discussions in general went well. The video sketches are a good means to initiate the discussion. As the group discussions took place at the end of the training day, some training groups were already really tired. The track trainers said that in these cases it was difficult to fill up the time available for discussion and therefore the group discussions were sometimes concluded earlier. We have to take into account that some track trainers didn't have much or any experience in carrying out group discussions with young people. It is likely that a more

experienced and better trained course leader would have no problem motivating the group, despite it being the end of an already long day and the participants were a little tired.

Some trainers hadn't entirely understood the procedure for the group discussion. They thought it was important to show and discuss each of the video sketches. In those cases, the trainers had to break off the conversation and ultimately there was little to no interaction between participants.

Planning and organization of the training days
The training days sometimes progressed in a rather chaotic manner. Because of the bad weather conditions, participants arrived too late or not at all. And participants who had agreed to use their own cars were ultimately afraid to use them. In these cases, the training centres had to provide one. According to the trainers, the participants were badly informed about the contents of the training programme. A lot of participants expected they would get a skill-oriented track training. This led to problems, especially in the case of Lelystad. In Lelystad, the regular training programme of the ANWB was taking place next to the NovEV training. Quite a few participants were disappointed that they could not have the regular (more spectacular?) ANWB track training. The ANWB trainers said it was difficult to keep these participants interested in the second phase training programme. In Rijssen, "regular" training also took place simultaneously, but this was not mentioned as a problem.

In the opinion of the trainers, most of the organizational problems could have been prevented if there had been someone co-ordinating all training activities on each training day. It would also have been helpful if all participants had done the training in hired cars. A 'hired car' or 'rented car' is a car that is provided by the project organisation. In the blueprint for the Dutch second phase training programme an important principle was that participants should do the training in their own car (or the car they are using most, in most cases being the car of one of the parents). This goes for the feedback drive as well as for the track training. Some of the participants couldn't come to the training with their own car, so we arranged a rented car for them (a car from the training institute or a car from a driving instructor). This complicated planning and organisation quite a lot. In addition, due to the bad weather conditions participants that came to the training or the driving audit in their own car ultimately refused to drive in their own car or weren't allowed by their parents (because it was their car). In these cases, considerable improvisation was necessary to provide a car for the participants. If we had worked with hired cars for everyone from the outset, none of these planning or organisational problems would have occurred.

2.4.3 Results of the discussion with the researchers

Recruitment of the participants
Looking back at the recruitment of the participants by telephone, it is possible that the participants had been given a too positive picture of the project and the training programme. In the first stage of recruitment, all young drivers were told that they could participate in a (spectacular) skid training day. In most cases it was not mentioned that the participants had to come back a month after the training for a second feedback drive and that they had to fill out several instruments (questionnaires and diaries). This may have caused a lot of participants to quit the project when they found out they were not able to do a skid training (but a safety training instead), they had to be present on two days and they had to fill out a questionnaire and keep a diary twice.

It probably would have been useful to ask participants to confirm their participation, and all the details of the project in writing, once they had committed themselves during the initial telephone recruitment.

When it comes to incentives for participation, a fee of €25 and a raffle with the chance of winning one of two travel vouchers or free car insurance for a year for free, the conclusion must

be that these incentives were not attractive enough for young people. Perhaps a more personal approach, in which young drivers are recruited by their driving instructor or by their examiner, would have been more effective.

Website, transfer of information, and data collection
The website had two main goals: transfer of information to participants and collection of data by Internet. As far as data collection is concerned, the website proved to be of great value. Most young people have access to the Internet and filling out questionnaires and diaries using the website was very efficient.
Transfer of information through the Internet was less successful. Sometimes relevant information was available too late (for instance, the description of the route to the training centres). The forum function of the website (encouraging discussion groups online) also didn't work as planned. There was not enough time to provide new information on a regular basis and to stimulate the participants to discuss issues with each other on the forum site.

2.4.4 Feedback form: Participants

After the final part of the project, namely the second feedback drive (one month after the training day), the young drivers all filled out a questionnaire in which they could indicate how satisfied they were with the training and the project. The most important purpose of this questionnaire was to find out what the young drivers (thought they had) learned during the training.

One of the questions was how "Fun" and "Useful" the participants thought the different parts of the project had been. Table 4 shows the percentages of respondent who "highly agreed" and "agreed" with the statements that the different parts were Fun and Useful.

Table 4 Percentage of respondents "highly agree" or "agree"

		Males (n=82)	Females (n=58)	Overall
Pre training feedback drive	Fun?	94	92	93
	Useful?	85	86	85
Training on track	Fun?	92	92	92
	Useful?	85	87	86
Group Discussion	Fun?	55	65	59
	Useful?	67	65	66
Post training feedback drive	Fun?	90	93	91
	Useful?	91	97	93

Table 5 shows that the respondents appreciated both feedback drives (before and after the training) the most. The young drivers were the least satisfied with the Group Discussion.

Because the training took place at different locations and therefore with different instructors, Table 5 shows the percentages (highly) agree with "Fun" and "Useful" at the different locations.

Table 5 Percentage of respondents "highly agree" or "agree"

		Experimental		Control (n=34)	Overall
		Lelystad (n=34)	Rijssen (n=62)		
Pre training feedback drive	Fun?	94	94	90	93
	Useful?	88	84	83	85
Training on track	Fun?	82	98		92
	Useful?	79	90		86
Group Discussion	Fun?	47	67		59
	Useful?	44	79		66
Post training feedback drive	Fun?	94	86	97	91
	Useful?	94	94	91	93

A remarkable result from Table 5 is that the control group seemed to appreciate both feedback drives almost the same as the experimental group. Usually, this is a problem for an experimental-control group design. The control group is often far less motivated, because they received a comprised version of what was promised. In this study, there does not seem to be a problem with the motivation of the control group.
Table 53shows that, overall, participants in Rijssen were more content with the training and discussion. There was not much difference in the assessment of both feedback drives. Participants in Lelystad found both feedback drives slightly more fun and useful, but this could be because they were not so content with the training on the track and the group discussion, therefore appreciating the feedback drives even more. For both locations, the group discussion was seen as the least attractive module in terms of "fun" and "usefulness". However, when the two locations are analysed separately it becomes clear that there are differences on all modules, and that the modules of the training day (group discussion and track training) are rated higher in Rijssen than in Lelystad. This indicates that the Rijssen group was more content than the Lelystad group. This phenomenon is the strongest in the group discussion. Here we saw a 35-percentage point difference on the "usefulness" of the group discussion. It would appear that, according to the participants, in Rijssen the group discussion succeeded in getting a "message across", in contrast to the group discussion in Lelystad.

It was also important for participants to receive the right message. In order to verify if this was the case, the questionnaire contained several statements.

Figure **19** shows the percentages of '(highly) agree' and 'disagree' with several statements. The percentages in the Figure represent the answers of both control and experimental groups, because there was no difference whatsoever between those groups in what they thought they had learned from the training. Promisingly, but not surprisingly, most respondents disagreed with the statements that they should look for more challenges and should drive in a more sporty manner. What is also important is that they do not think they drive better than the average motorist. Less than 10% of the participants agreed with this statement. It would have been a very undesirable side effect of the training, if the young drivers thought that they improved that much. However, only 40% of the participants believe they perform less well than the average driver. This must mean that more than half of the participants believe they are at least as good a driver as the average driver, with a lot more experience.

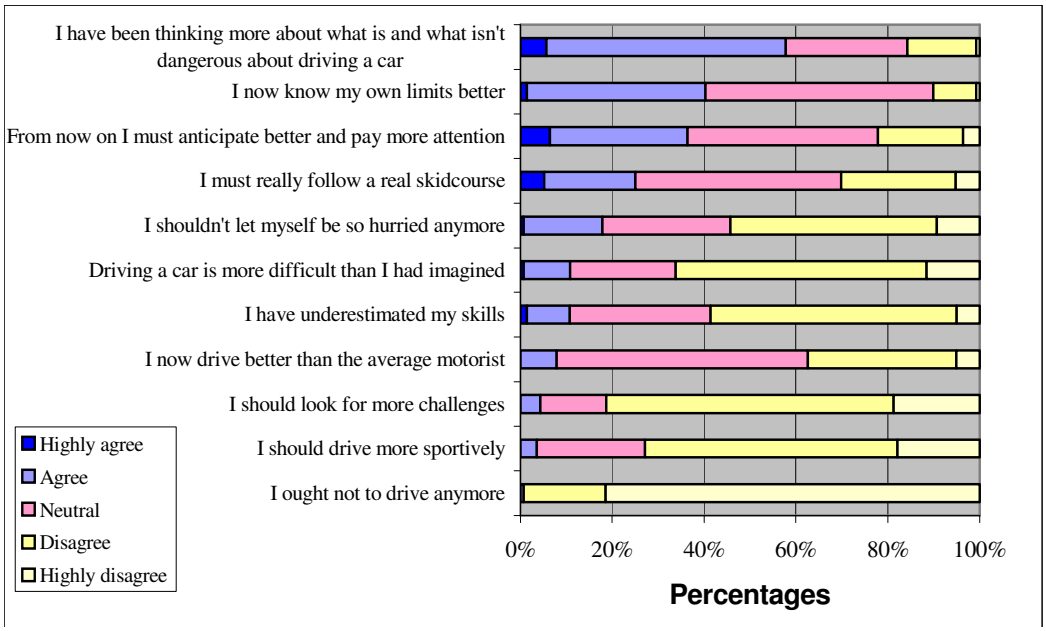


Figure 19 Percentages Agree – Disagree on what was learned during training

2.5 Evaluation design and timetable

As described in the introduction, the effectiveness of the second stage is dependent on several factors. In this evaluation we will not only study the ultimate effects, but also the process by which the effects come about. The latter is of importance to understand what elements need improvement in order to strengthen the effects.

In the evaluation study the following questions were addressed:

- Which changes (in knowledge, attitudes, intended behaviour, and driving behaviour) can be observed that can be attributed to the course;
- Does the training have an effect on those young drivers that are most at risk?
- How attractive is the training? (This topic has already been dealt with in par. 7.5.4)

Originally it was the intention to address the following questions as well. However, for several reasons, it was not possible to answer these questions in this report:

- What is the relationship between the quality of the basic training and the effectiveness of the second phase;
  Because of the number of participants dropping out of the programme, there were not enough participants to reach sufficient power in the statistical analysis for a distinction between different qualities of the basic training.
- What is the time span of the effect, and has the training led to new behaviour or insight that still develop after training;
  It would have been very interesting to see if the positive changes as a result of the training are still visible after a longer period, or if new developments occur after the training. However, it was not possible within the scope of this project to follow the participants for a longer time, (for example with a second post-test after a year).

2.5.1 Evaluation Design

The effect of the track training and group discussion was studied using a before-and-after design with a control group. Participants were randomly assigned to the control or the experimental (treatment) group. The *control group* (n=28) participated in both feedback drives. In addition to the feedback drive, the *experimental group* (n=99) also participated in the track training and the group discussion.

Table 6 shows a comparison between the experimental group and control group on four variables. Two other groups are also described. These are the 'no intervention' group, which consists of participants who were originally assigned to the experimental or control group but did not show up for the training day. Therefore they received no intervention at all, neither feedback drive, track training nor group discussion. The fourth group are those participants who did not show up for the second feedback drive. There is only information available on their performance before the training.

The comparison shows that the experimental group and control group do not differ on these four variables.

Table 6 Comparisons between experimental group and control group

		Experimental group (N=99)	Control group (N=28)	No intervention (N=15)	Drop out (N=36)
Gender	% Male	61%	61%	40%	56%
Age	Mean	20	20	21	20
Number of months drivers' licence	Mean	8	9	8	9
Hours of training for drivers' licence	Mean	38	41	46	40

Variables studied in the evaluation were: driving performance, risk awareness, self-assessment, and course satisfaction.

For research purposes, the most favourable design would be "double-blind". This would exclude shifts in results due to expectations of the participants and the assessor (the instructor). In a double-blind test design, neither the participant *nor* the assessor would know whether the participant is a member of the control or the experimental group and whether the feedback drive is before or after the training. In this study, the instructors were very much aware if the feedback drive concerned the first or second drive, for the simple fact that it was also their first or second series of feedback drives. The instructors were, however, not aware of the assignment to groups. It was not possible, of course, to prevent the participants and instructors from talking about their experiences in the programme so far.

2.5.2 Timetable and data collection

Table 7 Timetable

	December 2003 **Pre-test** One month before training day	January 2004 **Training day**	February 2004 **Post-test** One month after training day
Experimental group	Questionnaire Diary	Pre-test feedback drive - On-road observation form - Driving assessment Track exercises Group discussion	Post-test feedback drive - On-road observation form - Driving assessment - Satisfaction questionnaire
Control group	Questionnaire Diary	Feed back drive - On-road observation form - Driving assessment	Feed back drive - On-road observation form - Driving assessment - Satisfaction questionnaire

2.5.3 Estimated power of the design

The original design was to conduct the study with a sample of 300 young, inexperienced drivers. A distinction was made in the research design (Table 8) between drivers who attended a regular driving education and drivers who attended a 'best practice' education to obtain their drivers licence (RIS = Rijopleiding in Stappen = Stepwise Driver Training). The programme of such a 'best practice' education is very structured and a candidate can only pass from one module to another if all the training objectives of the former module are fully met. As a didactic principle candidates first have to learn so called handling scripts (this is the traffic situation, I want to do this (e.g. turn to the left) so I must first do this (e.g. look in the mirror) and then do that). What is also different compared with the regular driver training is that a four hour track training is included. This is not a short skid course. The intention is to let the candidate feel how easy it is to lose control and that is better to avoid certain situations than to rely on your skills.

Table 8 Intended research design

	Experimental group (training course)	Control (no course)
Regular driving education	100	50
'Best practice' education	100	50

Due to many respondents dropping out before and during the project, only 127 respondents finished the last part of the project. Before conducting any analysis, a power estimation was conducted to assess if there was a reasonable chance of finding any effects with these numbers of respondents. The question whether the remaining group of respondents was still representative, and not affected by a selection bias, has already been addressed.

Table 9 Actual research design

	Group		Total
	Experimental	Control	
Regular driving education	60	15	75
'Best practice' education	39	12	51
Unknown	0	1	1
Total	99	28	127

The question is how this modification affects the power of the experiment. In short, the power of a statistical test is the chance of finding a significant difference, if one is there. The power is dependent on three factors (Stevens, 1996):

1. The significance level (**α**) set by the experimenter
2. Sample size (**n**)
3. Effect size (**d**) – How much of a difference the treatments make, or the extent to which the groups differ in the population on the dependent variable(s).

For the estimation of the power in this experiment, the assumption was made that there is a moderate effect of the training on the dependent variables. The significance level (α) was set to .05. The sample size that was used for the power estimation was the smallest group comparison needed (the number of respondents with a 'best practice' education). This results in the following power estimations for a two-sided t-test, F-test, and Chi-square tests (Cohen, 1988).

Table 10 Power estimations

	t-test (α =.05; d=.50)	F-test (α =.05; f=.25; df=1)	Chi-square tests (α =.05; W=.30; df=1)
Original design (n=150; n=75)	.70	.86	.95
Actual design (n= 50; n=25)	.41	.42	.56
Actual design without 'best practice' vs. Regular (n=127; n=63)	.79	.80	.92

As can be seen in the table, the modification of the design does result in a decrease in power. Stevens (1996) argues that a study with a power of .70 or .80 is a good investment of money and resources. Therefore a comparison between the drivers with a 'best practice' and a regular education does not seem feasible. Without such a comparison, the power estimations are quite promising.

A problem may also arise when other subdivisions are made in the groups. When for example the gender of the drivers is inserted as a factor, the groups are divided in half. Not only the loss of subjects in this experiment has decreased the power of the study; also the introduction of extra (sub) groups (e.g. gender) in the design may lead to loss of power. In other words, we have too few subjects in each (sub) group, so the chance of finding a statistically significant difference is greatly reduced, even if such a difference actually exists.

2.6 Data collection methods

The following instruments were used in the evaluation:

- Questionnaires: about driving skills, self-assessment and risk awareness. This questionnaire was based on previous work by Hatakka (1998). In this questionnaire, risk awareness was also measured using photographs of "normal" traffic situations.

- Diaries: Semi-structured questionnaires in which driving events were reported by the participants. The results of these diaries exceed the scope of this report, and will not be presented here.

- On-road observation form: an assessment tool to describe the driving performance of a driver. The driver himself, the passenger, and the driving instructor completed these forms after the feedback drive. This instrument was not only used for research purposes. By comparing the three assessments, the forms were also used as input for the discussion after the feedback drive.

- Driving assessment: assessment by the instructor of the quality of driving in three fields: vehicle control, driving skills, and calibration skills.

- Satisfaction questionnaire: this questionnaire contains questions on how satisfied participants were about the different components of the training day and the feedback drives. In addition, they were asked about what they thought they had learned from the training.

2.6.1 Website

For the administration of the instruments in the Pre-test and Post-test period (questionnaire and diary) a website was used. This website was only accessible for participants. This allowed for a day-to-day overview of those who responded. The participants that did not respond were encouraged by e-mails and phone calls to do so.

3. Evaluation Results

The effect of the training was measured with several instruments. The results from each of these instruments will be discussed in the following sections. This chapter begins with the results from the questionnaire which the participants filled out before and after the training. Secondly, the Driving assessment form will be discussed. This is a form, which was filled in by the instructors (privately) after the feedback drives before and after the training. Finally we will discuss the direct calibration scores that were given by the instructors, but also by the participants themselves, about their performance in the feedback drive.

3.1 Questionnaire

The questionnaire (see annex 10) was filled out by the participants approximately one month before and one month after the training. The participants were invited by email to visit the website and fill in the questionnaire. The participants who did not have access to the internet received the questionnaire by mail.

This questionnaire focused on self-assessment of driving skills and risk, and safety awareness. The questionnaire consisted of four parts:

1. General questions (Age, Gender, How often do you drive?, etc.)
2. Items focusing on self assessment: weak and strong skills
3. Items focusing on risk and safety awareness: the degree of difficulty and complexity of the driving and traffic task
4. Judgement of traffic situations

3.1.1 Self-assessment and risk awareness

Figure 20 shows the participants' opinions on several statements concerning risk awareness. A distinction has been made between risks caused by the young, novice driver himself and risks caused by other road users. The higher the percentage in Figure 20, the less young drivers are concerned with that item, indicating low risk awareness. So, on the left are subjects that concern the young drivers very much, and on the right are subjects that the young drivers are not concerned so much about. An alarming result is the fact that the young drivers do not seem to be concerned about driving too fast. A least 60% of the respondents are not concerned about driving too fast.

For some of the risk factors, the respondents had to indicate the perceived risk of the same situation twice. Once the situation where the young drivers were the cause (internal risks) and once when other road users caused the risk (external risks). There seems to be a slight difference in estimation of internal and external risks. For example, young drivers are more concerned when other drivers take risks in traffic than when they take risks themselves. A Principal

Component Analysis was conducted to investigate if the internal and external caused risks were two different constructs. The results indicated this was not the case. Young novice drivers generally assess risks as less or more dangerous, irrespective of the cause of that risk.

The percentages indicating risk awareness are plotted for several groups, experimental-control; before and after the training. A third group is plotted in the Figure, the 'no intervention' group (before and after training). This group consists of respondents who did not show up for the first feedback drive or training day, but did fill in their questionnaire twice (pre- and post-'training'). This is a special group because there is information in the before and after situations, but they did not have any intervention whatsoever. This means that any development within this group can be attributed to the passing of time.

The hypothesis is that the experimental – after group, which should have been affected by the training, scores a lot lower on most of the risk awareness items. However, there does not seem to be a pattern of differences between groups in the Figure, for risks caused by the driver himself or by other road users.

A Repeated Measures Analysis was conducted on the questionnaire items concerning risk awareness. This analysis compared the scores before and after the training (within components) of the experimental group with the control group (between components). No effect of the training was found.

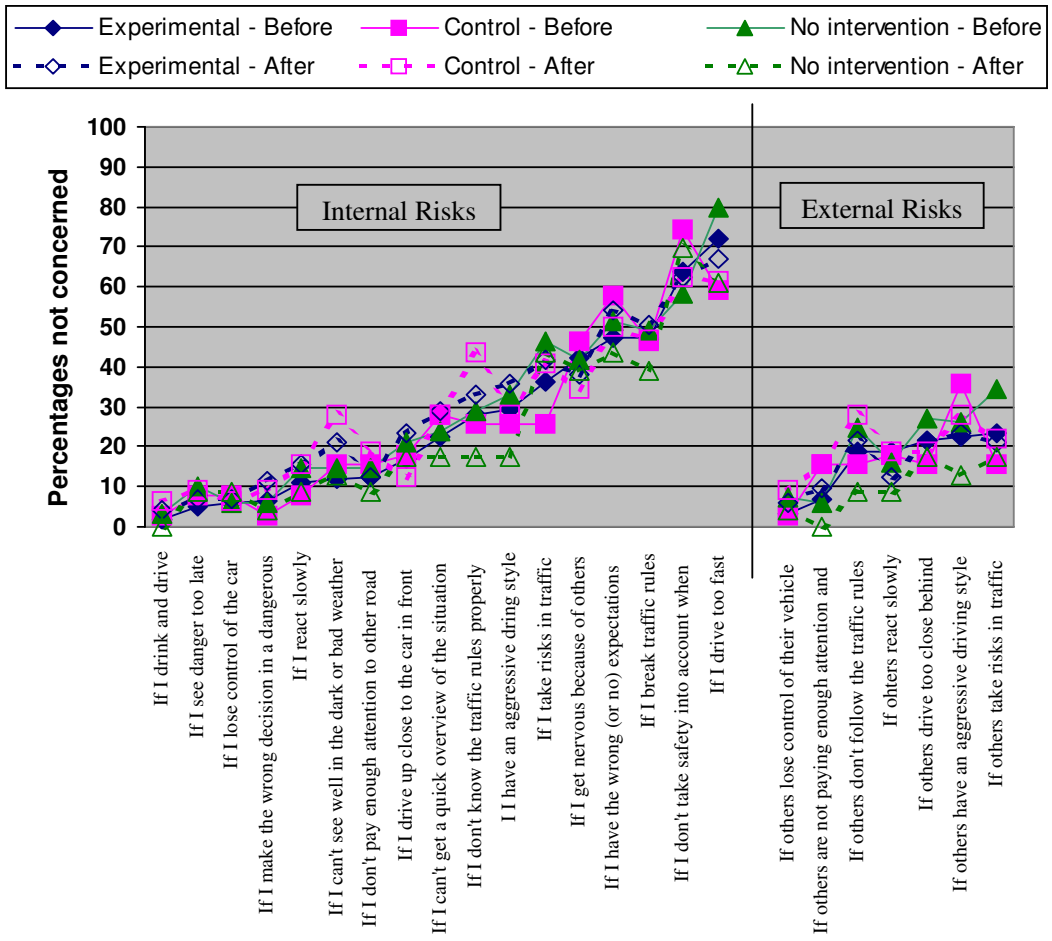


Figure 20 Risk Awareness (grouped into internal and external risk factors) – Percentage not concerned about issue

Figure 21 shows the scores on the self-assessment of driving skill items. A distinction was made between skills involving "safe driving" (e.g. obeying traffic rules) and "demanding driving" (e.g. reacting accurate to emergency situations).The higher the percentage, the more participants think they are (very) strong in that aspect, indicating a possible problem of overestimation of skills. So, on the left are driving skills where the participants think they are not so strong; on the right, are aspects of driving where the young drivers think they are (very) strong.

There seem to be some 'outliers' in the Figure, especially in the "No intervention-After" group. This is a result (gimmick) of the way the data was presented, and does not actually indicate extreme answers. The items were ordered on the basis of the percentages of the "Experimental-Before" group, because this was the largest group. The outliers merely indicate that those groups have a different ordering of aspects they think they are (very) strong in. They do not, overall, have a bigger problem with overestimation of skill..

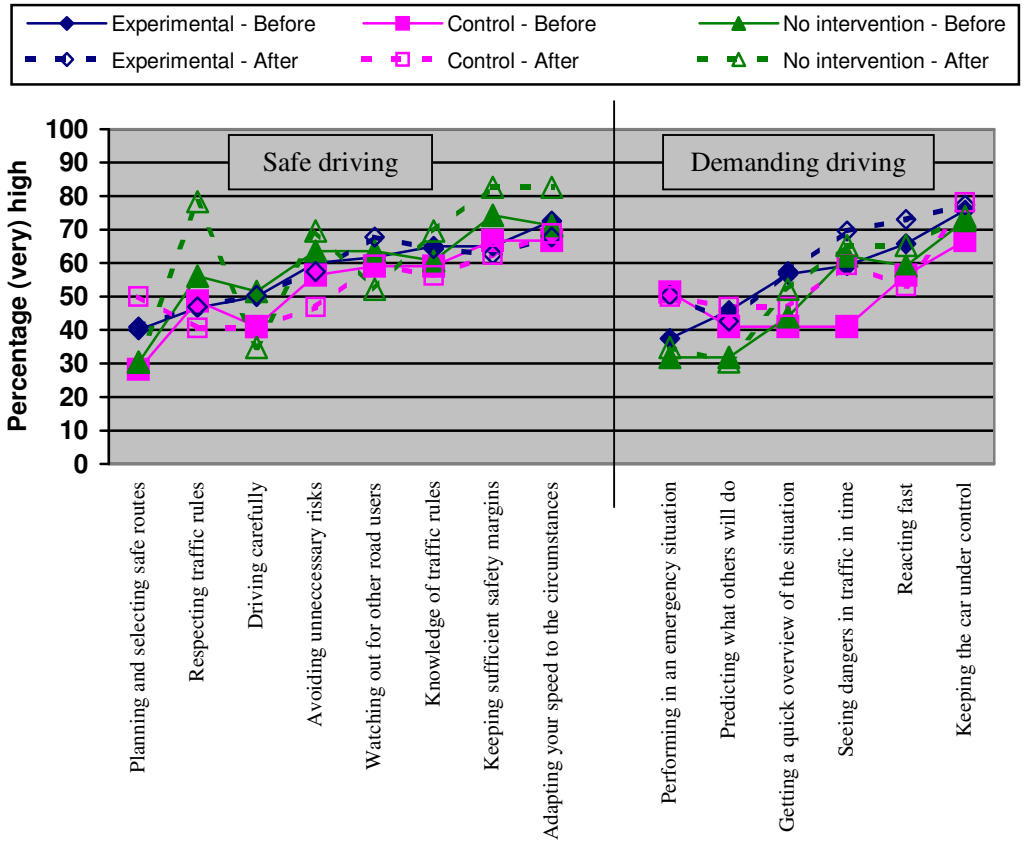


Figure 21 Self-assessment of driving skill (grouped into 'safe driving' skills and 'demanding driving' skills) -- Percentage (very) strong

Figure 21 shows that young drivers are, overall, rather confident about their driving skills. At least 30% of the participants believe they are (very) strong in all skills, and in some skills even more than 60% believe they are (very) strong. There is no difference in skills concerning 'safe driving' and risks concerning 'demanding driving'. A Principal Component Analysis also indicated that young novice drivers generally do not make a difference between 'safe driving' and 'demanding driving' when assessing their skills.

A Repeated Measures Analysis was conducted on the questionnaire items concerning self-assessment of driving skill. This analysis compared the scores before and after the training

(within components) of the experimental group with the control group (between components). No effect of the training was found.

Gender Differences

Further investigation was conducted to see if there were effects of the training that were not so obvious. For example the Repeated Measures Analysis was conducted for Females and Males separately. A Multivariate Variance Analysis (MANOVA) indicated that there is a significant main effect of gender on risk awareness and self-assessment (Table 11 and 12 show the Pillai's Trace statistics). Males score worse on risk awareness, in the sense that they see less risks. Males also score worse on self-assessment, in the sense that they are more confident than females.

Table 11 MANOVA (b) - Risk awareness

		Value	F	Hypothesis df	Error df	Sig.	Partial Eta^2
Intercept	Pillai's Trace	,999	12985,660[a]	24,000	367,00	,000	,999
Gender	Pillai's Trace	,115	1,995[a]	24,000	367,00	,004	,115

a Exact statistic
b Design: Intercept + Gender

Table 12 MANOVA (b) - Self-assessment of skill

		Value	F	Hypothesis df	Error df	Sig.	Partial Eta^2
Intercept	Pillai's Trace	,984	1641,133[a]	14,000	380,00	,000	,984
Gender	Pillai's Trace	,320	12,754[a]	14,000	380,00	,000	,320

a Exact statistic
b Design: Intercept + Gender

Within these groups of females and males, no effects of the training were found. That is, there was no effect of the training found for either the males or females. In the power estimation, it was concluded that there would not be enough participants to include an extra factor (gender) and still have a reasonable chance at significant results. It is, however, unlikely that this was the reason for not having any significant results in this analysis, because the effect size was also quite small.

The fact that an effect of gender on the several statements was found, indicates that the questionnaire does measure something. It is possible that the questionnaire measures more stable personality traits, such as those that can be expected to differ between males and females, (in other words, not something that can be changed with a one-day training course).

Driver education

The analyses were also repeated for the participants with a regular driver education and for participants with a 'best practice' driver education. No effects of the training were found within these groups. In section 2.5.3, the conclusion was reached that the chance of finding significant effects of the training with the distinction between regular and 'best practice' education would be very slim. So this result was not very surprising.

However, the power should not be a problem for finding the main effects of driver education. But there was no main effect of the driver education found. The participants with a regular driver education filled in the questionnaire in the same way as the participants with a 'best practice' education.

3.1.2 Calibration

Calibration is defined as the balance between self-assessment of skill and risk awareness. Those who see little or no danger in traffic, but are at the same time highly confident of their own skills form a risk group in terms of calibration. In order to analyse if the training had any effect on calibration, as was measured by the questionnaire, the group of participants was split in half, twice. First the group was split in half on the basis of their assessment of their driving skill. Participants who had a lower score than the median formed one group; they perceive themselves with the weakest skills. Participants with scores above the median formed another group; they perceive themselves with the strongest skills.
Secondly, the group was split in half, in the same manner, on the basis of their perception of risks in traffic. This resulted in Table 13, for the experimental group.

The risk group, in terms of calibration, is highlighted. These are the participants who express the most confidence about their own driving skill and, at the same time, see little danger in traffic. The expectation for the training is that the number of participants in this group is smaller after the training. But there is another group which indicates bad calibration (also highlighted). This is the group of drivers who are very insecure of their own driving skill and, at the same time, see much danger in traffic. This is not a risk group, in itself, but it is important that this group does not get bigger as a result of the training. It is not the intention of the training to create young drivers who are too insecure of their driving skills and who see too much danger in traffic.

Unfortunately, as can be seen in Table 14, there was no shift in the experimental group from 'bad' calibration to 'good' calibration. There is even some increase in the number of participants who see no danger in traffic and assess themselves to be very skilful. This is however not enough for a significant effect of the training.

Also in the control group (Table 14) there was no distinct shift in calibration groups.

Table 13 Calibration development - Experimental group

Pre training		Self assessment of skill		
		Low	High	
Risk awareness	No danger	28	17	45
	Much danger	20	32	52
		48	49	97

Post training		Self assessment of skill		
		Low	High	
Risk awareness	No danger	26	20	46
	Much danger	21	30	51
		47	50	97

Table 14 Calibration development - Control group

Pre training		Self assessment of skill		
		Weak	Strong	
Risk awareness	No danger	7	2	9
	Much danger	10	9	19
		17	11	28

Post training		Self assessment of skill		
		Weak	Strong	
Risk awareness	No danger	10	4	14
	Much danger	6	8	14
		16	12	28

Table 15 shows the effect of the training on calibration for the individual participant. The rows in the table denote the 'calibration-group' of the participant before the training and the columns denote the 'calibration-group' after the training. This means for example that of the 60

participants who were in the 'good-calibration' groups before the training, 40 remained in that group, 12 ended up in the 'insecure-group' and 8 in the 'risk-group'.

Table 15 indicates that there is some change between groups before and after the training, but there is no specific pattern. Some participants improve, whereas others end up in a worse group than before the training. The same effect can be seen in the control group.

Table 15 Shifts in Calibration

		Experimental group				Control group			
		After training				After training			
		Good	Insecure	Risk	Total	Good	Insecure	Risk	Total
Before training	Good	**40**	12	8	60	**12**	1	3	16
	Insecure	11	**7**	2	20	5	**5**	0	10
	Risk	5	2	**10**	17	1	0	**1**	2
	Total	56	21	20	97	18	6	4	28

3.1.3 Situation questions

Besides the questions concerning "Risk awareness" and "Self-Assessment of skills", the respondents were asked to judge traffic situations on photos. The respondents were asked to estimate how fast they would drive through the situation displayed on the photo. See annex 10 for the two situations that were used. Both situations were shown in pairs, randomly scattered throughout the questionnaire. In one of the photos, the situation was slightly more complicated (because of the presence of a cyclist). The objective of these questions was to measure if youngsters take the difficulty of the situation into account when judging how to handle the situation. In this case, it was not important that they responded with low speeds, as such, but it was important that they adapted their speed to the complexity of the situation.

These situations were tested in a pilot study (n=10), by means of the website. During a discussion afterwards, it turned out that none of the young drivers in the pilot study noticed the small differences in the photos. There were even some complaints that the website did not work correctly because the same photos seemed to be shown several times. However, it turned out that the young drivers subconsciously did alter their speed, dependent on the complexity of the situation.

Figure 22 shows the percentage of 'good' responses for the different groups concerning the first situation. A 'good' response means that the person reported a lower speed when the situation was more complex. A 'bad' answer was considered to be the case when there was no difference in speed between the two situations, or when the speed was even higher in the more difficult situation. The Figure shows the responses of (1) the Experimental group, which attended a feedback drive, training on a track, and a group discussion; (2) the Control group, which only attended the feedback drive; but also (3) an extra group – No intervention – these are participants who did not show up for the first day, but did complete the questionnaire before the second feedback drive. The 'no intervention' group did not get 'anything' between the before-training and after-training moments.

As can be seen in the Figure, there were no effects found in the training on the judgement of the situation.

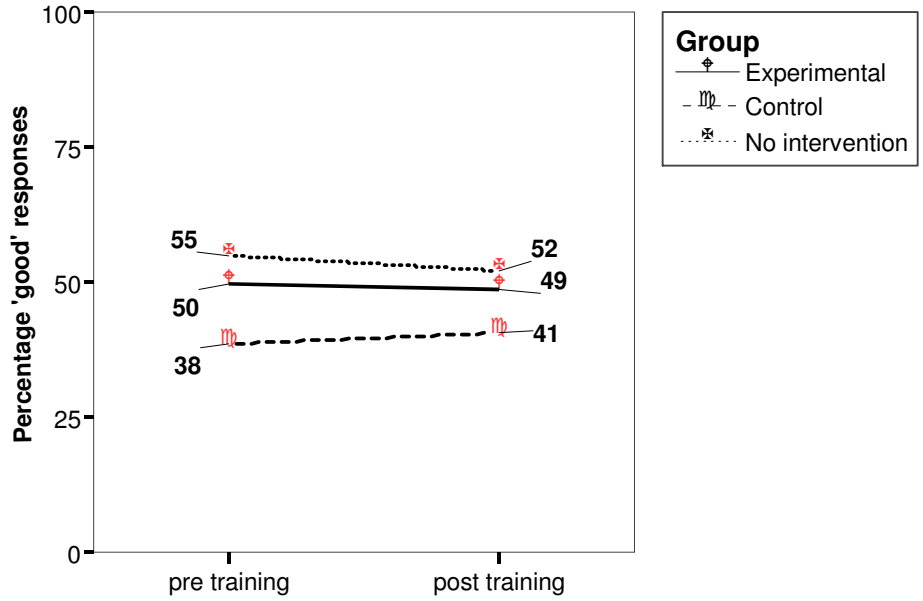


Figure 22 Percentage 'good' responses in Situation 1

Annex 10 includes the second situation that was displayed on photo, after which the participants were asked to estimate the speed they would drive. Figure 23 shows the percentage of good responses by the participants from the experimental and control groups, before and after the training. A 'good' answer means that the respondent reported a lower speed when the situation was more complex.

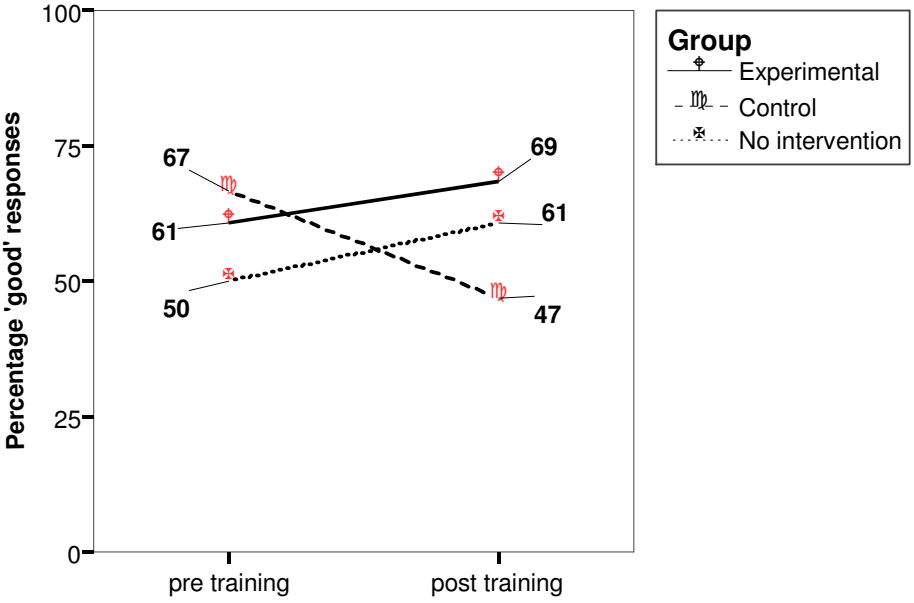


Figure 23 Percentage of 'good' responses in Situation 2

Figure 23 shows a (slight) improvement in the experimental group after the training. The control group diminishes substantially at the same time, which can not be explained so easily. Another puzzle is the improvement in the 'no intervention' group, which is about the same as in the experimental group.

Table 16 shows the exact number of participants with a 'right' or 'wrong' response to the situations for experimental vs. control group, before and after the training.

Table 16 Frequencies of right and wrong responses to situation 2

	Experiment	Control	No intervention	Total
Pre training				
Right response	73	26	33	132
Wrong response	47	13	33	93
Total	120	39	66	225
Post training				
Right response	79	15	14	108
Wrong response	36	17	9	62
Total	115	32	23	170

Using Log linear analysis several models were tested for significance to find the factors which could explain the data (Table 65). The factors involved were [A] right or wrong response, [B] before or after the training, and [C] Experimental, control or no intervention group. A model without the training- group interaction did not fit the data (significant deviation). A slightly more complicated model, including the training interaction, did fit the data sufficiently. In order to test if the interaction factor was the cause of the model fitting the data, the difference between the chi-square and degrees of freedom between the models was calculated. This resulted in a significant contribution of the training – group interaction in the fit of the model.

Table 17 Log linear analysis Situation 2

Model		Chi-square	Df	Significance
Without Training interaction	[AB][AC]	18.456	4	< 0.0001
With Training interaction	[AB][AC][BC]	4.828	2	n.s.
Training Significance	[BC]	13.628	2	< 0.01

Of course, these analyses still do not explain why the improvement that was found in the experimental group also took place in the group that did not receive any form of intervention, nor even a feedback drive. Nor does it explain why the control group, which should have remained at the same level, showed such deterioration. Both effects could have been caused by the small number of participants in the control and no intervention group. The log linear analysis was repeated for the control and no intervention scores added up (because both groups represented 'no training'). There was still a significant effect of the training which can be attributed to the training.

3.1.4 Conclusions questionnaire

The results from the questionnaire are somewhat unclear: some effects of the training were found, but they were not consistent, and were not always in the expected direction.

An alarming result from the questions concerning risk awareness is the fact that the young drivers do not seem particularly concerned in general, and especially not about driving too fast. A least 60% of the respondents do not believe that it is dangerous to drive too fast. On the other hand, it turned out that young drivers are, overall, rather confident about their driving skills. At least 30% of the participants believe they are (very) strong in all skills, and in some skills more than 60% believe they are (very) strong. Unfortunately, these opinions did not improve after the

training. The fact that there were significant gender differences in these issues, led to the conclusion that this part of the questionnaire probably measured more stable attitudes or personality traits (which could not be changed with a one-day training course or within the period of a month).

There were some significant effects in the situation judgements, but they were somewhat contradictory. There was only an effect in one out of two situations. However, the control group, which should theoretically remain at the same level, scored lower on the post-test. Moreover, there is no explanation why the group without any intervention improved in the same way the experimental group did.

3.2 Driving Assessment

For every on-road feedback drive, an assessment form was completed by the instructor (see annex 10). The participants were judged on twenty skills (items), according to the scale: insufficient (1 point), sufficient (2 points), or good (3 points). The twenty skills can be divided into three groups, Vehicle control, Driving skills, and Calibration skills. Three new variables were formed by adding up all the scores in each group. So, for example, for vehicle control, the total score of a participant could range from 3 (insufficient on all three items) to 9 (good on all three items). The internal consistency, Cronbach's alpha, is depicted after each scale.

Vehicle control (Range: 3 – 9; Cronbach's alpha: **0.67**)

1. Preparation for drive / end of drive
2. Vehicle handling
3. Vehicle control

Driving skills (Range: 8 – 24; Cronbach's alpha **0.71**)

13. Driving on straight and bendy roads
14. Behaviour at junctions
15. Behaviour when turning
16. Entering and exiting traffic
17. Overtaking and passing
18. Being overtaken
19. Changing lanes and lateral positioning
20. Driving on different surfaces

Calibration skills (Range: 6 – 18; Cronbach's alpha: **0.84**)

7. Defensive behaviour: anticipation
8. Defensive behaviour: effective observation skills
9. Defensive behaviour: safety cushion and safety margins
10. Risk awareness, danger recognition, and traffic insight
11. Adapted and decisive driving: speed
12. Adapted and decisive driving: decisive handling

Analysis showed there is a significant correlation (ranging from .55 to .65) between these three aspects of the driving assessment ($p<.000$). Participants who score high on, for example driving skill, also score high on vehicle control.

Because the training course was aimed at improving calibration, the largest effect was expected in the calibration score. Note that this calibration (measured with the feedback drive) is not exactly the same as calibration measured by the questionnaire).
Repeated Measures Analysis on each skill was used to test this hypothesis.

3.2.1 Vehicle control

The total score on vehicle control for the test and control group, before and after the training, is shown in Figure 24. The vertical axis depicts the mean value of the total score on vehicle control. The range of this total score was from 3 (insufficient on all three items) to 9 (good on all three items). Because the instructors rarely used the 'insufficient' category, the range in Figure 24 has been modified from 6 (sufficient on all three items) to 9 (good on all three items). The horizontal axis depicts the value on the first and second feedback drive.

Two extra groups are displayed. The "No intervention group" indicates the participants who could not come to the first feedback drive, but were still invited to attend the second. This group did not receive any form of training or even feedback. The "drop out group" consists of respondents who dropped out of the project after the first feedback drive, so we do have information on their performance during the first feedback drive. There is only information on their driving skill at one moment. The whiskers shown in Figure 24 display the 95% confident interval.

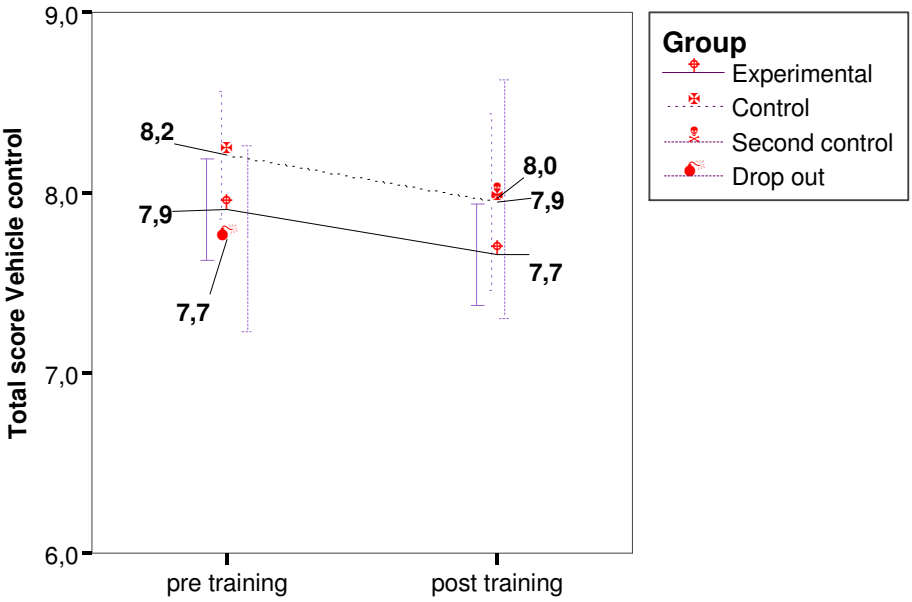


Figure 24 Vehicle control (total score from driving assessment)

As can be seen, both groups (control and experimental) perform slightly worse on driving skill after the second feedback training. What is remarkable is that there seems to be a consistent difference between experimental and control group. Although respondents were assigned randomly to the conditions, the control group performed (slightly) better on the first drive.

A Repeated Measures Analysis was conducted to discover if the differences between before and after training (and differences between control and experimental group) were significant. Although all participants of either one or two feedback drives are shown, the analysis was conducted only on the young drivers participating on both days. The respondents who did not show up for the first (no intervention group) or second feedback drive (drop out) were deleted from the analysis.

No significant effects were found. Also, the effect sizes of these factors were negligible.

There were some indications that the training that took place in Lelystad differed somewhat from the training in Rijssen, for example, in appreciation by the respondents. Therefore, the

results were also analysed for Lelystad and Rijssen as a factor (see Figure 25). The Repeated Measures Analysis revealed no significant effects.

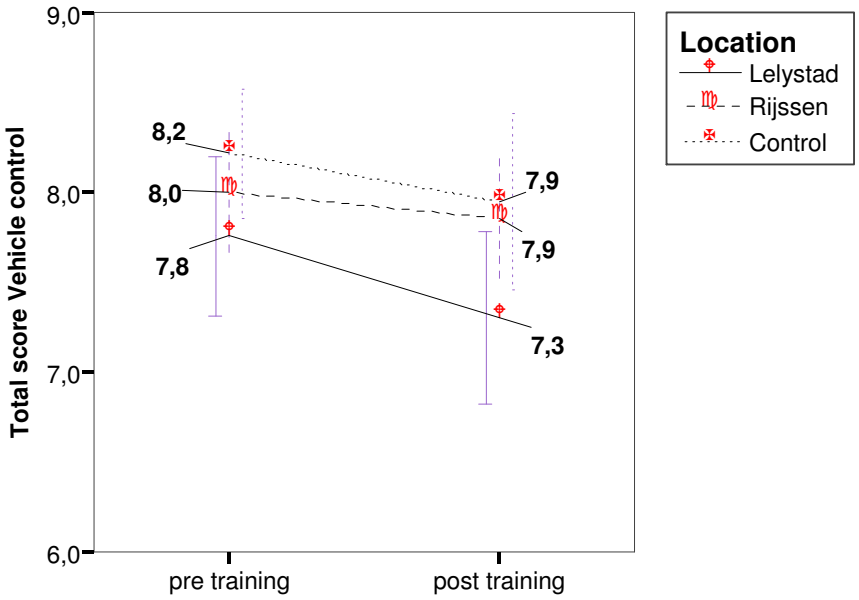


Figure 25 Vehicle control (Lelystad and Rijssen separate)

These results (no significant effects of the training) were just as expected. Because the training was not designed to affect vehicle control performance, it is not surprising that no effect was found.

3.2.2 Driving skill

Figure 26 shows the total score on driving skill, before and after the training The vertical axis depicts the mean value of the total score on driving skill. The range of the total score was from 8 (insufficient on all eight items) to 24 (good on all eight items). Because the instructors rarely used the 'insufficient' category, the range has been modified from 16 (sufficient on all eight items) to 24 (good on all eight items). The horizontal axis depicts the value on the first and second feedback drive.

No effect of the training is expected, because the training was not aimed at improving driving skill. Of course, it is not a negative result if there are some effects on driving skills, as long as they are in the desired direction.

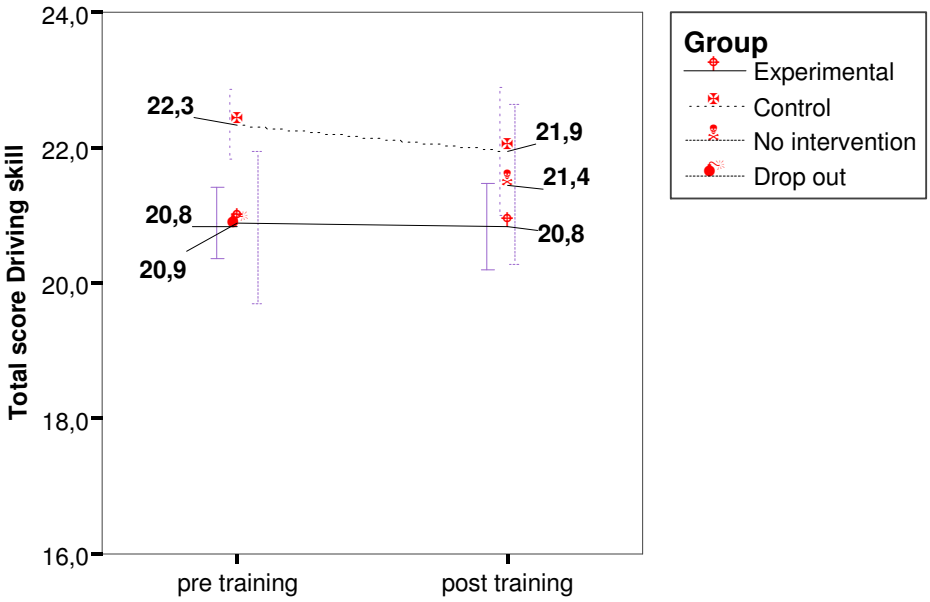


Figure 26 Driving skill (total score from driving assessment)

Again these results were analysed with Repeated Measures Analysis, and a significant effect of "Group" was found between the subjects in the analysis (Table 67). In other words, there was a significant difference in performance (before training) between experimental and control group. This is remarkable, because respondents were assigned randomly to the conditions, so there should not be a difference between control group and experimental group during the first feedback drive. An explanation for this phenomenon could be the following: the instructors in the experimental conditions had to come to an unfamiliar location. They indicated themselves that it had been a complicating factor to carry out the feedback drives in unknown surroundings. The instructor who performed the feedback drives for the control group did this in a familiar location. There was also a difference between the locations themselves. The experimental locations were rather remote. The control locations, on the other hand, were located in the middle of a city.

This difference between the control group and experimental group has some consequences for the interpretation of the results. But this will be addressed later, because in driving skill there were no effects found as a result of the training (Table 18).

Table 18 Tests of Between-Subjects Effects – Driving skill

		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Source	Intercept	1266,515	1	1266,515	10534,650	,000	,988
	Group	1,122	1	1,122	9,335	,003	,070
	Error	14,908	124	,120			

Table 19 Multivariate tests (b)– Driving skill

		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
training	Pillai's Trace	,000	,031(a)	1,000	124,000	,860	,000
training * Group	Pillai's Trace	,001	,151(a)	1,000	124,000	,699	,001

a Exact statistic
b Design: Intercept+Group Within Subjects Design: training

Figure 27 shows the experimental group split up in both locations, Lelystad and Rijssen.

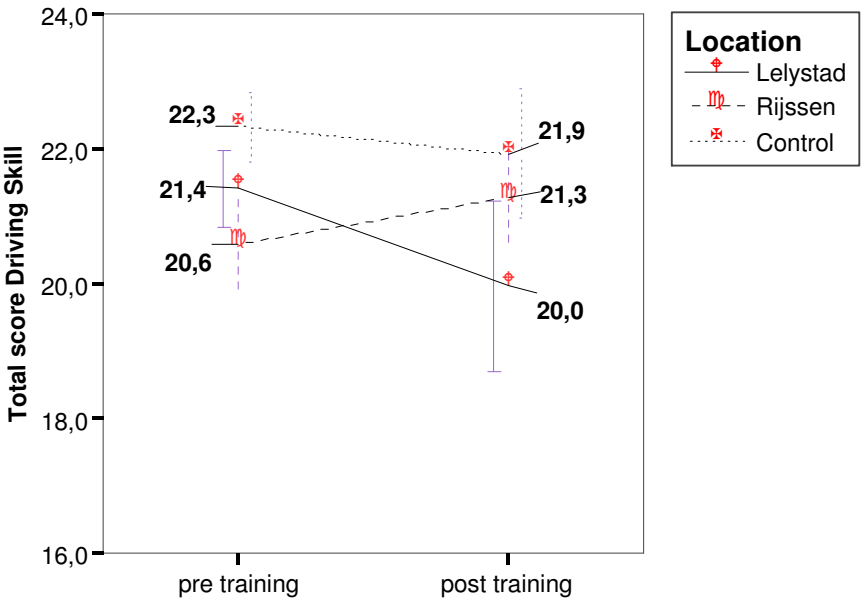


Figure 27 Driving skill (Lelystad and Rijssen separate)

This Figure shows a clear difference between Rijssen and Lelystad. The performance of the young drivers who participated in the training in Lelystad shows roughly the same decline as the control group. On the other hand, the Rijssen participants improved their performance in a statistically significant manner. Repeated Measures Analysis indicated that the effect of the training on the different locations was significant (Table 20). The Partial Eta Squared indicates the effect size of the training. An Eta Squared of .055 can be seen as a moderate effect size (Stevens, 1996).

Table 20 Multivariate tests (b)– Driving skill (Lelystad and Rijssen separate)

		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
training	Pillai's Trace	,003	,385(a)	1,000	123,000	,536	,003
training * location	Pillai's Trace	,055	3,604(a)	2,000	123,000	,030	,055

a Exact statistic
b Design: Intercept+location Within Subjects Design: training

3.2.3 Calibration

Because the training was specifically aimed at improving calibration, the expectation is that the calibration score from the driving assessment was affected the most. Figure 28 shows this total score on calibration skill, before and after the training. The vertical axis depicts the mean value of the total score on calibration skill. The range of this total score was from 6 (insufficient on all six items) to 18 (good on all six items). Because the instructors rarely used the 'insufficient' category, the range in Figure 28 has been modified from 12 (sufficient on all six items) to 18 (good on all six items). The horizontal axis depicts the value on the first and second feedback drive.

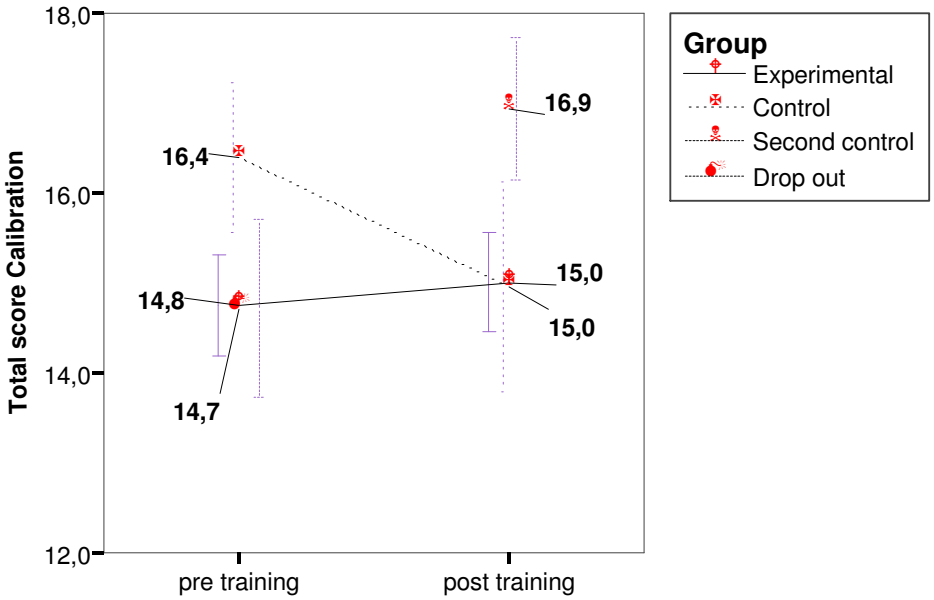


Figure 28 Calibration score (total score from driving assessment)

This Figure shows an interaction effect between Group and Training, which turned out to be significant (Table 22). However, the experimental group shows almost no improvement. The interaction is actually caused by the deterioration in the control group, after they had been assessed remarkably highly during the first feedback drive. We already saw the same significant difference between the experimental and control group in driving skill. This leads to the tentative conclusion that the control group and experimental group were assessed differently during the first feedback drive. The interaction effect in calibration score can be explained fully by this phenomenon. The interaction effect can be reduced to a different assessment during the first feedback drive. At the time of the second feedback drive, the difference was removed and the groups performed identically.

Also noticeable in Figure 28 is the outstanding performance of the 'No intervention' group. These are participants who could not come to the first feedback drive, but were still invited for the second. In the post-training situation they drove for the very first time. It's inexplicable how it is possible that this group without any training, or even a feedback drive, performed so well. A cautious guess is that these young drivers were more impressed and therefore more motivated than the rest of the group. Either that, or they were judged more leniently by the instructors, because it was their first time.

Table 21 Multivariate tests (b) – Calibration score

		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
training	Pillai's Trace	,018	2,234(a)	1,000	124,000	,138	,018
training * Group	Pillai's Trace	,034	4,309(a)	1,000	124,000	,040	,034

a Exact statistic
b Design: Intercept+Group Within Subjects Design: training

Figure 29 shows the performance of Lelystad and Rijssen.

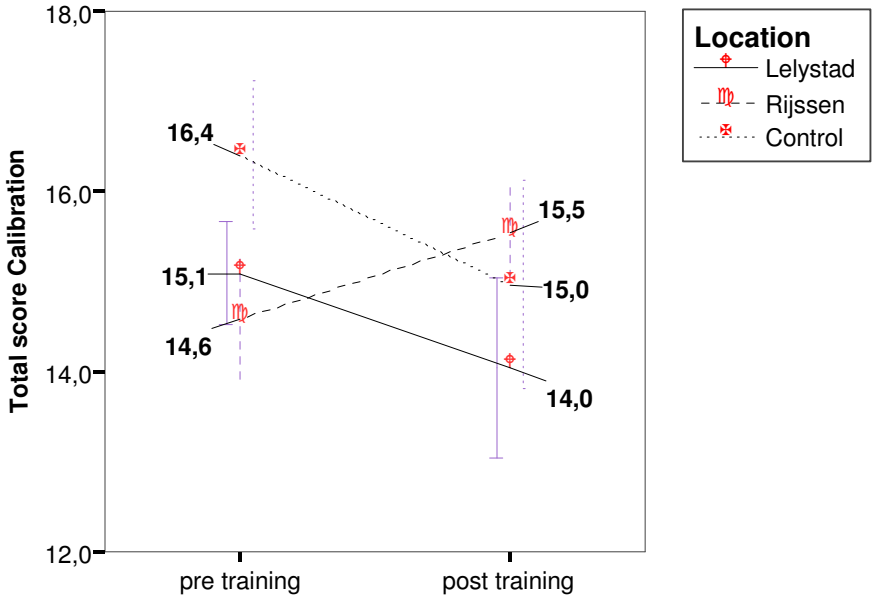


Figure 29 Calibration score (Lelystad and Rijssen separate)

Again there is a difference in performance between Rijssen and Lelystad. The young drivers who went to Lelystad perform roughly the same as the control group who did not receive a training and group discussion. The participants from Rijssen show a totally different development: they improve from the first to the second feedback drive, in a statistically significant manner. What is important in this comparison is the fact that there was no difference between the performance of Lelystad and Rijssen during the first feedback drive. The effect of the training (Table 22 indicates that this effect is significant) can be attributed completely to the training.

Table 22 Multivariate tests (b)– Calibration score (Lelystad and Rijssen separate)

		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
training	Pillai's Trace	,017	2,187(a)	1,000	123,000	,142	,017
training * location	Pillai's Trace	,083	5,559(a)	2,000	123,000	,005	,083

a Exact statistic
b Design: Intercept+location Within Subjects Design: training

3.2.4 Conclusions Driving Assessment

The results on driving assessment show that there is a significant effect of the training on driving and calibration skill. There is a difference between experimental and control group. Because there is a significant difference between these groups before the training, this result is difficult to interpret.

However, there is also a significant difference between the performance in both locations, Lelystad and Rijssen. Because the performance in Lelystad and Rijssen was the same before the training, there are no problems with interpreting this result. The participants from Rijssen perform significantly better than the participants form Lelystad, as a result of the training.

In all measures (vehicle control, driving skill, and calibration skill) there seems to be a difference between experimental and control group (significant for driving skill and calibration skill), before training. This could be a problem, for example in Figure 28, where the effect of the training could be explained by the differences between control group and experimental group during the first driving assessment. The respondents were assigned to the conditions randomly, so the groups should be the same. Another explanation for this effect could be that the young drivers in the experimental group were assessed by different instructors than the drivers in the control group, or more likely, the different type of location of the experimental and control group. Because of the organization of the training day, the experimental feedback drive was in the vicinity of the training site (so not in a familiar environment for participants), which was a rather secluded area, while the control group feedback drive was in the vicinity of the participant's hometown and in a more urban area. When comparing the performance of Lelystad and Rijssen, these analytical problems are no longer an issue. The performance of the participants in Lelystad or Rijssen are exactly the same during the pre-training drive, and only differ on the post-training drive.

An effect that is apparent from all measures is that only the respondents who received their training in Rijssen perform better in the second feedback drive than in the first (with the exception of vehicle control were no group improves their performance). The respondents in the control group and from Lelystad perform even worse than before. How is this possible? The respondents in the control group should, in theory, remain at the same level. One of the explanations could be the fact that during the first feedback drive there was extremely bad weather, snowstorms, slippery roads, etc. This group of young drivers had never experienced these conditions. It is possible that they were so in awe of the extreme conditions that they all accepted very large safety margins, therefore scoring better on the feedback drive with regard to safe driving. During the second feedback drive the circumstances were normal and it is possible that they were assessed worse. But, if this is the case, all scores during the first feedback drive should have been lower. Unfortunately we can only speculate, and will never know what the scores would have been if the weather had been normal.

3.3 On-road observation form

The on-road observation form consisted of 3 parts and was to be completed by the instructor as well as the participant after the feedback drive. During the first feedback drive, another participant who was in the car also filled in the observation form. The opinion of this extra passenger has, however, not been analysed because there was no passenger present during the second feedback drive.

The first three items on the observation form (part 1), concerning assessment of driving skill, were to be rated on a scale from ‘weak’ (1) to ‘strong’ (5). The fourth item (part 2), concerning perception of the driving task, was to be rated on a scale from ‘difficult’ (1) to ‘(too) easy’ (5).

The final item (part 3) concerned the so-called 'driver profile'. The instructor or the participant could choose one out of five profiles that felt most appropriate. These profiles included:

(1) Very insecure driver who sees danger everywhere and allows for too large safety margins.
(2) Calm driver, very aware of the dangers of driving.
(3) Smooth driver, but sees the dangers of driving and the limits of his own abilities.
(4) Sharp driver who leaves himself limited room for manoeuvre, but has an eye for dangerous situations.
(5) Overconfident driver who leaves insufficient (safety) margins.

As previously mentioned, the training took place at two locations: Lelystad and Rijssen. In the previous section it has been shown that the two locations of the experimental group did not behave in the same fashion. Therefore, in this section, all analyses are performed comparing 'locations' (Lelystad, Rijssen, and Control) rather than comparing 'groups' (Experimental and Control).

In the following section, the results of several analyses are presented per item on the form. Furthermore, the results are divided into three subsections. The first subsection describes the scores provided by the instructors and compares the scores given before and after training. The second subsection describes the scores given by the participants and also compares before and after training scores. The third subsection compares the scores provided by the instructors to scores provided by the participants and considers the relationship between these scores before and after training.

3.3.1 Scores provided by the instructors

To analyse the effect of training on the scores provided by the instructors (for participants in each location) Repeated Measures Analyses were performed. The results for each item on the on-road observation form are presented and discussed below.

Vehicle control and general skills
The results, which can be found in Table 23, show a significant difference in scores before and after training over the different locations. Also, an Eta Squared of .094 could be considered to be moderate to large (Stevens, 1996), which is an indication of the effect size of training.

Table 23 Multivariate tests (b) – Vehicle control and general skills

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Training	Pillai's Trace	,022	2,698(a)	1,000	118,000	,103	,022
training * location	Pillai's Trace	,094	6,126(a)	2,000	118,000	,003	,094

a Exact statistic
b Design: Intercept+location Within Subjects Design: training

Figure 30 shows the scores on vehicle control and general skills, before and after the training. The mean score is depicted on the vertical axis, where (1) indicates 'weak' skills and (5) indicates 'strong' skills. The whiskers shown in Figure 30 illustrate the 95% confidence intervals.

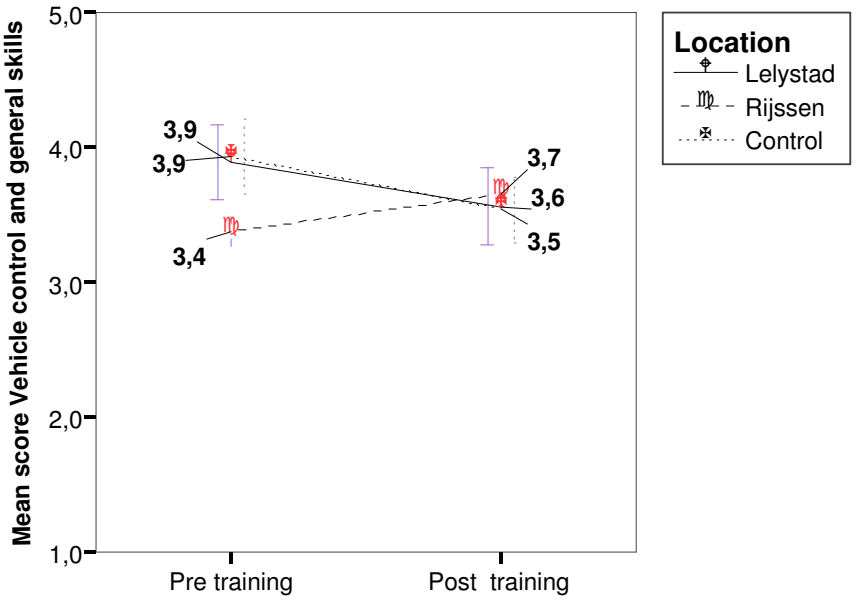


Figure 30 Vehicle control and general skills

Figure 30 displays a difference in particular between Rijssen versus Lelystad and the Control group. It shows that participants in Rijssen received a lower score from the instructor before training than participants in Lelystad and the Control group. This is actually contrary to expectation, as all groups are expected to start at a comparable level. Analyses in the previous sections already discovered a difference between the experimental groups and control group. However, in those analyses it was the control group that scored differently on the first feedback drive. In the current analysis it is the Rijssen location that shows a deviation.

It also shows that, according to the instructor, only participants in Rijssen were stronger in controlling the vehicle after training than before. Participants in Lelystad and the Control group were considered to be weaker in controlling the vehicle after training than before.

Safe and defensive driving

Figure 31 shows the scores on safe and defensive driving, before and after the training. The mean score is depicted on the vertical axis, where (1) indicates 'weak' skills and (5) indicates 'strong' skills.

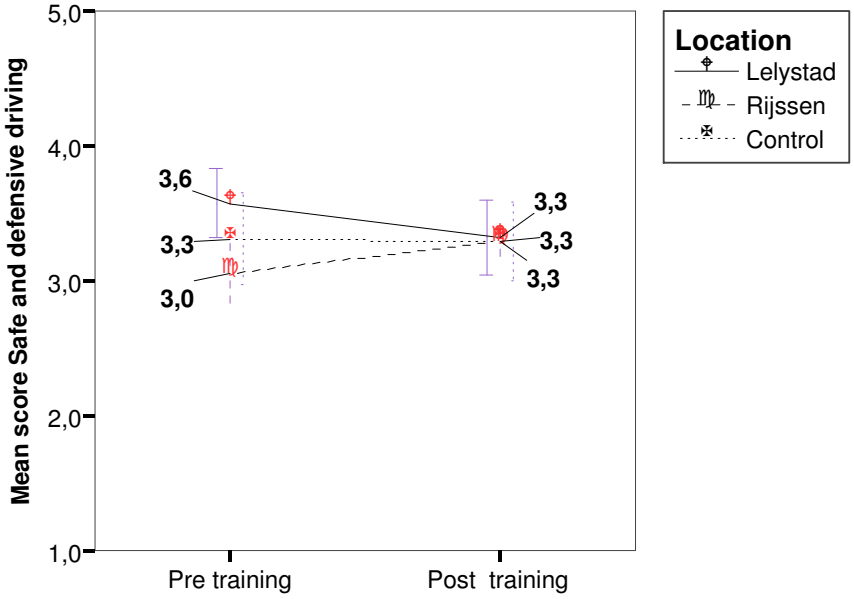


Figure 31 Safe and defensive driving

The results for the 'safe and defensive driving' item, show no significant differences in scores before and after training, either dependent or independent of location. As also found for the first item on the on-road observation form, Figure 31 shows that, according to the instructors, the participants in Rijssen were weaker before training in safe and defensive driving than participants in Lelystad or the Control group. Another similarity with the first item is that only participants in Rijssen received higher scores after training than before training.

Anticipation

The results, for the 'anticipation' item, which can be found in Table 24, show a significant difference in scores before and after training over the different locations. Eta Squared amounted to .208 and could be considered to be rather large, which is an indication of the effect size of training.

Table 24 Multivariate tests (b)– Anticipation

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
training	Pillai's Trace	,009	1,133(a)	1,000	119,000	,289	,009
training * location	Pillai's Trace	,208	15,653(a)	2,000	119,000	,000	,208

a Exact statistic

b Design: Intercept+location Within Subjects Design: training

Figure 32 shows the scores on anticipation, before and after the training. The mean score is depicted on the vertical axis, where (1) indicates 'weak' skills and (5) indicates 'strong' skills.

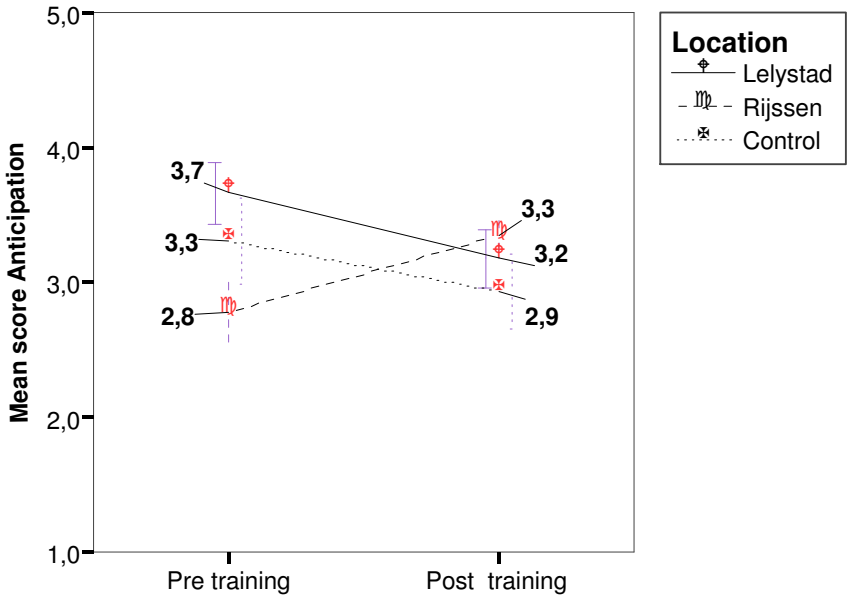


Figure 32 Anticipation

Similar to the results for the previous two items on the on-road observation form, Figure 32 shows that, according to the instructors, the participants in Rijssen were weaker before training in anticipation than participants in Lelystad or the Control group. Also, an opposite effect is found for the participants in Rijssen versus the participants in Lelystad and the Control group. Participants in Rijssen are considered by the instructor to be stronger in anticipation after training than before training. Participants in Lelystad and the Control group are given lower scores for anticipation after training than before training.

Estimation of the complexity of the driving task

The results, which can be found in Table 25, show a significant difference in scores before and after training independent of the location. Eta Squared amounted to .066 and could be considered to be moderate, which is an indication of the effect size of training.

Table 25 Multivariate tests (b)- Estimation of the complexity of the driving task

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
training	Pillai's Trace	,066	8,095(a)	1,000	114,000	,005	,066
training * location	Pillai's Trace	,026	1,515(a)	2,000	114,000	,224	,026

a Exact statistic
b Design: Intercept+location Within Subjects Design: training

Figure 33 shows the scores on estimation of the complexity of the driving task, before and after the training. The mean score is depicted on the vertical axis, where (1) indicates that the participant perceives the driving task as difficult, and (5) indicates that the participant perceives the driving task as (too) easy. This means that a high score is unfavourable. (In the previous Figures, a high score represented a favourable score).

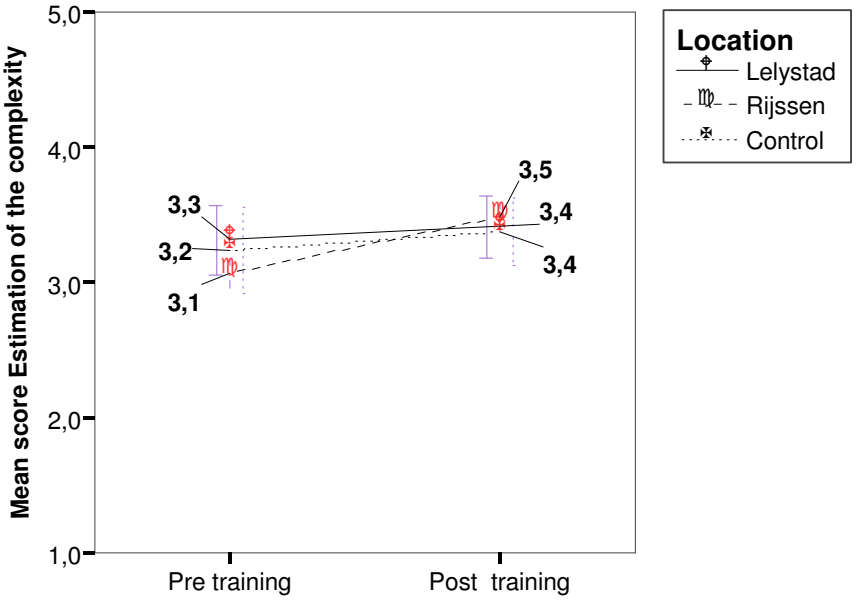


Figure 33 Estimation of the complexity of the driving task

This item on the on-road observation form shows a somewhat different result compared to the first three items. Figure 33 displays that the instructor rated the participants from all locations to estimate the complexity of the driving task as harder before the training compared to after the training. Or, in other words, according to the instructor, during the second feedback drive, the participants underestimated the driving task more than during the first feedback drive,. Because this effect is found for all locations, this implies an effect of the feedback drive, and / or maybe even an effect of the instructors (interpreting this item differently the second time).
Most probably this effect is caused by the extreme weather conditions during the first feedback drive. On some training days, a traffic warning was even issued not to go outside unless absolutely necessary. So it is unlikely, with such extreme conditions, that the participants underestimated the driving task.

Profile scores
The scale belonging to the item 'driver profile ' on the on-road observation form has a somewhat different scale than the previous four items. Namely, this item does not vary from 'weak' to 'strong' or from 'difficult' to 'easy'. In this case, one could assert that the scale varies from 'undesirable' to 'desirable' (the central profile) and back to 'undesirable' again. As this does not constitute an interval scale, it did not seem appropriate to conduct a Repeated Measures Analysis. Thus, frequencies and percentages were computed to detect differences (see Table 26 and Figure 34).

The results show that independent of location or effect of training, the central 'optimal' profile is chosen most. The most salient effect of training seems to occur around profile 3 and 4, which is, however, largely due to the smaller number of participants in Lelystad (45 vs. 23) and Rijssen (73 vs. 47) after training.

Table 26 Frequency and percentage for each profile before vs. after training

	Lelystad				Rijssen				Control			
	Pre training		Post training		Pre training		Post training		Pre training		Post training	
Profile 1	0	(0%)	0	(0%)	2	(3%)	2	(4%)	1	(4%)	0	(0%)
Profile 2	6	(13%)	5	(22%)	16	(22%)	13	(28%)	5	(21%)	8	(22%)
Profile 3	24	(53%)	13	(57%)	29	(40%)	24	(50%)	11	(46%)	23	(62%)
Profile 4	12	(27%)	5	(22%)	23	(32%)	6	(13%)	5	(21%)	5	(14%)
Profile 5	3	(7%)	0	(0%)	3	(4%)	2	(4%)	2	(8%)	1	(3%)

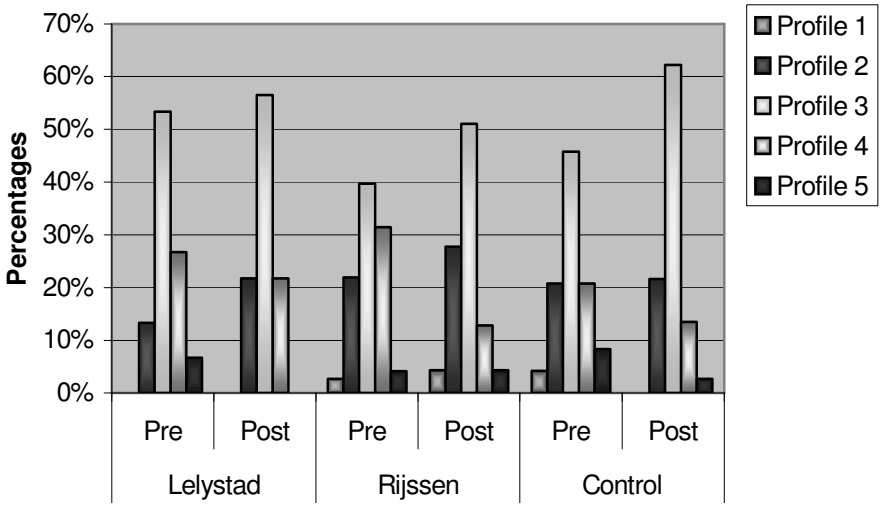


Figure 34 Percentage each profile is chosen by an instructor

Conclusions
The results presented in the above section, seem to lead to the conclusion that the training did have an effect on the scores given by the instructors. The effect of training is mainly found in the 'Rijssen' location and less in the 'Lelystad' location. However, to account for this effect is rather difficult, as a strange phenomenon seems to have occurred in Rijssen before the training. For all items, Rijssen starts out with the lowest scores compared to the other two locations before training. As mentioned earlier, this is contrary to expectation, as all locations should behave about the same before a manipulation is administered. In previous sections, a difference between the two experimental locations and the control group was also noted, for which there actually was a satisfactory explanation. But at this moment in time, no evident explanation comes to mind that can account for the deviation of Rijssen, although it seems that this effect should not be attributed to differences between the participants in the Rijssen location. A more plausible explanation would be that the instructors in Rijssen interpreted the items on the form in a different way than instructors in the other locations or than after training.

3.3.2 Scores awarded by the participants

To analyse the effect of training on the scores awarded by the participants in each location, Repeated Measures Analyses were performed. As none of the results were significant, only figures are shown.

Initially it was intended that the on-road observation form would be filled in, not only by the instructor, but also by the participant, following each feedback drive. However, due to miscommunication, not all participants completed the form. Especially the last item on the

form, the driver profile, was seldom filled in by the participant. Therefore, differences between driver profiles provided by the participant were not analysed.

Items 1 to 4
The results, which can be found in Figure 35, show no significant differences in scores before and after training.

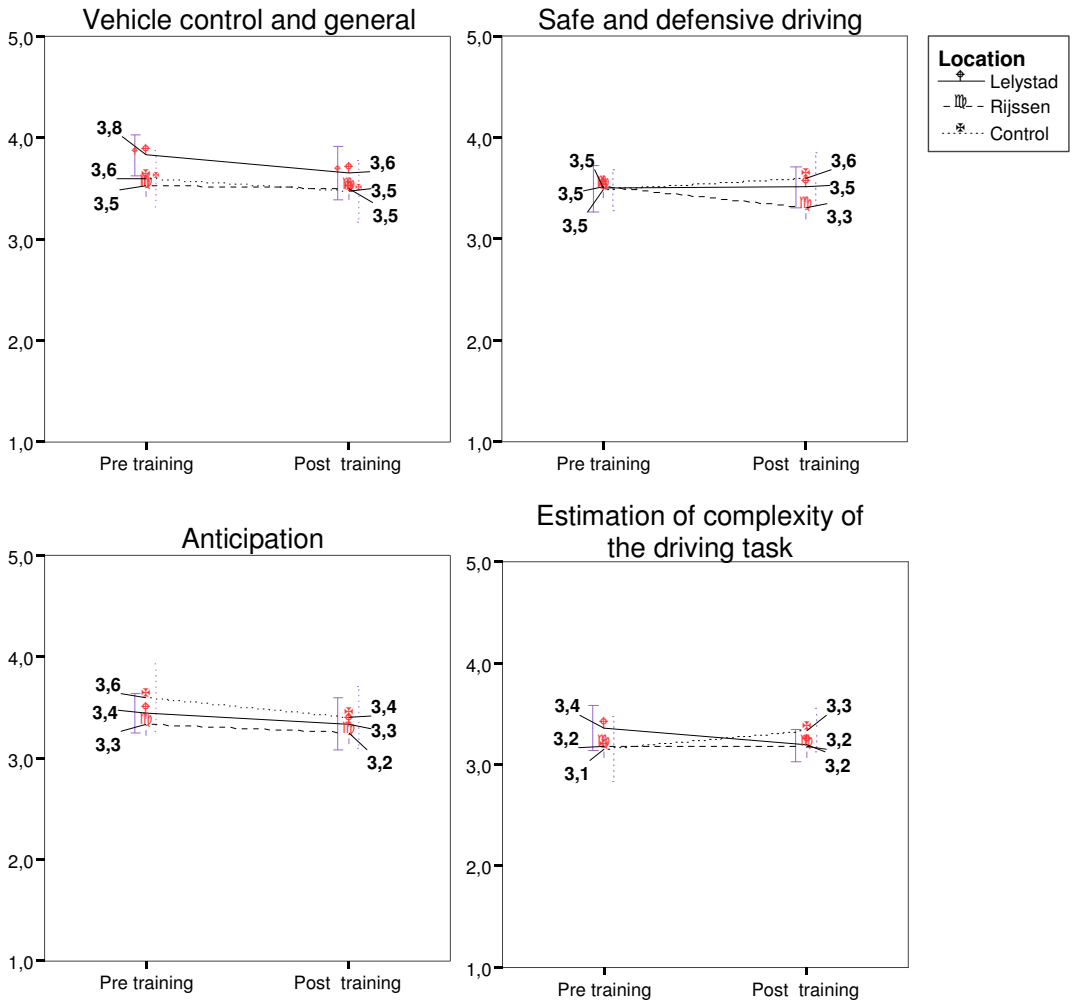


Figure 35 Results for the participant – Mean scores

Conclusions
An important objective of the training was to improve calibration. This would be the case if the participants did not see any improvement in their skills, after the training, but perceived the complexity of the driving task as being more difficult.

Repeated Measures Analyses were performed to detect significant differences between scores provided by the participants before and after training. However, none of the items on the on-road observation form filled out by the participant showed significant differences between before and after training, although an effect of training was expected. It seems that , according to the young drivers themselves, the training did not have an effect on the assessment of skill or complexity of the driving task. This is, on one hand, positive because the participants did not

think they improved their skills as a result of the training. On the other hand, there was no change in their perception of the complexity of the driving task.

3.3.3 Comparing the instructors with the participants

The first step in analysing the differences between the scores awarded by the instructors and scores given by the participants, was to calculate correlations between the scores. The Pearson correlation coefficients, although most of them significant, ranged between .13 and .43. Because of this weak but significant correlation, Repeated Measures Analyses were conducted to find more specific patterns in the comparison between instructors and participants. It is to be expected that before training, participants will overestimate or perhaps underestimate their skills more than after training. Therefore, two Repeated Measures Analyses were conducted per item. The first considering the scores awarded by both the instructor and the participant before training, and the second considering the scores given after training.

Vehicle control and general skills
Table 27 shows some results for both raters in all locations concerning the first item.

Table 27 Descriptive statistics – Vehicle control and general skills

	Condition	Pre training			Post training		
		Mean	Std. Deviation	N	Mean	Std. Deviation	N
Instructor	Lelystad	3,80	,795	44	3,58	,830	33
	Rijssen	3,36	,901	70	3,67	,769	61
	Control	3,89	,667	36	3,48	,671	42
	Total	3,61	,850	150	3,59	,755	136
Participant	Lelystad	3,82	,540	44	3,67	,777	33
	Rijssen	3,60	,646	70	3,49	,698	61
	Control	3,61	,645	36	3,48	,707	42
	Total	3,67	,620	150	3,53	,719	136

Pre training:
The results show no significant differences between instructors and participants before training. This finding is actually not what is expected, but is less surprising than if scores had differed significantly in the unexpected direction, (if the participants had rated themselves to be less skilful than the instructors rated them). The results, which can be found in Table 28, also show an effect of location, which is again unexpected and even undesirable before training.

Table 28 Multivariate tests (b)- Vehicle control and general skills (Pre training)

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rater	Pillai's Trace	,000	,003(a)	1,000	147,000	,958	,000
rater * location	Pillai's Trace	,051	3,947(a)	2,000	147,000	,021	,051

a Exact statistic
b Design: Intercept+location Within Subjects Design: rater

Post training:
The results again show no significant differences between instructors and participants. This is not very surprising as it would be expected that, after training, the difference between instructors and participants would not exist, or at least be less than before training. As in this

case, there was no difference found before training, an effect of training in the expected direction should not be found.

Safe and defensive driving

Pre training:

The results, contrary to the results for the first item, show a significant interaction effect between rater and location before training (Table 29). Eta Squared for this effect could be called moderate (.048). This effect was not expected, as all participants should be rated about the same before training in each location.

Table 29 Multivariate tests (b)- Safe and defensive driving (Pre training)

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rater	Pillai's Trace	,024	3,712(a)	1,000	148,000	,056	,024
rater * location	Pillai's Trace	,048	3,705(a)	2,000	148,000	,027	,048

a Exact statistic
b Design: Intercept+location Within Subjects Design: rater

Post training:

The results (Table 30) show a significant effect of rater independent of the location the participants were in. It would be expected that the difference in scores between instructor and participant would diminish after training. Therefore it seems hard to explain how the difference between raters can be significant after training, while it was not before training. Table 31 shows that the difference in mean score between the raters is actually less after training compared to before training and is therefore not an unexpected result. However it should be noted that this effect cannot be attributed to the training, as this result is found independent of location. This result could therefore be interpreted as an effect of the feedback drive.

Table 30 Multivariate tests (b)- Safe and defensive driving (Post training)

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rater	Pillai's Trace	,030	4,146(a)	1,000	132,000	,044	,030
rater * location	Pillai's Trace	,012	,811(a)	2,000	132,000	,446	,012

a Exact statistic
b Design: Intercept+location Within Subjects Design: rater

Table 31 shows some descriptive statistics for both raters in all locations concerning safe and defensive driving.

Table 31 Descriptive statistics – Safe and defensive driving

	Condition	Pre training			Post training		
		Mean	Std. Deviation	N	Mean	Std. Deviation	N
Instructor	Lelystad	3.39	0.841	44	3.30	0.810	33
	Rijssen	3.07	0.976	71	3.28	0.710	61
	Control	3.33	0.793	36	3.29	0.716	41
	Total	3.23	0.903	151	3.29	0.732	135
Participant	Lelystad	3.48	0.628	44	3.52	0.566	33
	Rijssen	3.55	0.789	71	3.31	0.564	61
	Control	3.28	0.701	36	3.54	0.596	41
	Total	3.46	0.728	151	3.43	0.580	135

Anticipation

Pre training:

The results, which can be found in Table 32, show a significant effect of rater independent of the location the participants were in. Eta Squared equals .065 and could be considered moderate. It would be expected that before training, participants tend to score themselves as 'stronger' at anticipation compared to instructors. Table 33 shows that this is indeed the case. Besides an effect of rater, an interaction effect between rater and location is also found (Table 33), which seems to be an undesirable effect. However, Table 33 shows that again, a strange phenomenon has occurred with the instructors in Rijssen before the training. Unfortunately, it is not yet clear what this phenomenon exactly is.

Table 32 Multivariate tests (b)- Anticipation (Pre training)

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rater	Pillai's Trace	,065	10,310(a)	1,000	148,000	,002	,065
rater * location	Pillai's Trace	,044	3,435(a)	2,000	148,000	,035	,044

a Exact statistic

b Design: Intercept+location Within Subjects Design: rater

Table 33 Descriptive statistics – Anticipation

	Condition	Pre training			Post training		
		Mean	Std. Deviation	N	Mean	Std. Deviation	N
Instructor	Lelystad	3,45	,791	44	3,15	,619	33
	Rijssen	2,89	1,008	71	3,34	,750	61
	Control	3,19	,786	36	3,00	,775	41
	Total	3,13	,926	151	3,19	,738	135
Participant	Lelystad	3,45	,589	44	3,33	,736	33
	Rijssen	3,42	,822	71	3,25	,745	61
	Control	3,53	,878	36	3,27	,807	41
	Total	3,46	,772	151	3,27	,757	135

Table 33 shows that before training, participants rate themselves stronger on anticipation than the instructors do, which is a result in accordance with expectation.

Post training:
The results show that the significant difference between raters that was found before training is not found after training. Again, as the interaction between rater and location is not found significant, this effect can be attributed to the feedback drive.

Estimation of complexity of the driving task
Pre training:
The results show no significant differences between instructors and participants before training. This finding is actually not what was expected, but less surprising than when scores would have differed significantly in the unexpected direction. Also, the results show no effect of location, which is to be expected before training.

Post training:
The results, which can be found in Table 34, show a significant difference between instructors and participants after training. Eta Squared is .060 and can be considered moderate. However, this effect is not found in combination with location, and therefore this effect cannot be attributed to training, but perhaps to the feedback drive.

Table 34 Multivariate tests (b)- Estimation of complexity of the driving task (Post training)

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rater	Pillai's Trace	,060	8,094(a)	1,000	126,000	,005	,060
rater * location	Pillai's Trace	,041	2,667(a)	2,000	126,000	,073	,041

a Exact statistic
b Design: Intercept+location Within Subjects Design: rater

Table 35 shows that, after training, participants consider the driving task to be easier than the instructor does, compared to before training, which is an unintended effect. Among other things, the training intended to make participants more aware of the complexity of the driving task. The results here, however, indicate that the opposite effect has been achieved.

Table 35 Descriptive statistics – Estimation of complexity of the driving task

	Condition	Pre training			Post training		
		Mean	Std. Deviation	N	Mean	Std. Deviation	N
Instructor	Lelystad	3,39	,722	44	3,41	,665	32
	Rijssen	3,13	,803	69	3,53	,570	57
	Control	3,24	,855	34	3,25	,630	40
	Total	3,23	,794	147	3,41	,620	129
Participant	Lelystad	3,34	,608	44	3,19	,471	32
	Rijssen	3,23	,789	69	3,18	,658	57
	Control	3,24	,781	34	3,25	,670	40
	Total	3,27	,734	147	3,20	,617	129

Conclusions
The results discussed above do not present strong evidence for an effect of the training, but should perhaps be interpreted as evidence for a positive effect of the feedback drive.

Overall, the hypothesis was that differences between instructors and participants would diminish as an effect of training. This is found in a few instances, but the results in previous sub-sections show that this is mostly due to a change in the instructors instead of a change in the participants. Therefore, based on the results of the on-road observation form, it does not seem that the training had the intended effect on the participants.

Generally, from these results, it can be concluded that while the instructors did see some improvement as a result of the training, the participants did not.

4. Discussion

4.1 The evaluation study: strengths and weaknesses

The evaluation was a true field study in the sense that not all conditions met the scientific criteria. Some "complications" were apparent from the start (like the familiarity with the traffic situations during the feedback drive), others happened just "by accident" (the trainers' opinions about the principles of the track exercises). Here, an overview is given of the strength and weaknesses of the study, as well as an analysis of the impact the "weaknesses" have had on the reliability, generalizability, and the validity of the results.

The selection of subjects/participants
Any study that aims to document the effects of a voluntary course like this one, needs to be able to reject the claim that the participants in the study were a very small and select group: namely persons that are very interested in safety issues. The study was able to document non-response both at the start and during the study. We had a good overview of certain critical variables about the characteristics of the 68% that refused to participate in the study. A comparison between the refusers, dropouts, and participants on four variables (gender, age, driving experience and hours of training) did not show major differences between the groups, which led to the conclusion that, based on these variables, the participants could still be regarded as an a-select group. However, they did of course differ in the most important variable: namely to accept the invitation to sign up for the course and/or to finish it.

The participants were randomly assigned to the two research conditions (control and experimental group). Only in a few cases this was not possible, such as in the case of scheduling problems, or the (im)possibility of a participant to come to one of the training sites.

The validity of the feedback drive
In studies of this nature, the actual driving behaviour is seldom assessed. In this study the driver was assessed during his pre-test feedback drive and the post-test feedback drive. The value of these assessments can be questioned on three aspects:

As it was part of the course, the instructor influenced the candidate by commenting on his behaviour. To make the conditions between control and experimental group comparable, he did this for both the control and the experimental group. The consequence is that the 'control' group experienced some form of intervention as well.

Because of the organization of the training day, the pre-test feedback drive was in the vicinity of the training site (so in an unfamiliar environment for the participant) while the post-test feedback drive was in the vicinity of the participant's hometown. For the control group both feedback drives were in the vicinity of their hometown. Taking these circumstances into account, it is to be expected that in comparison to the control group, the experimental group

would perform worse on the pre-test feedback drive. Moreover, as both groups took their post-test feedback drive in a familiar driving environment, the experimental group should improve in comparison to the pre-test drive, and the control group should neither improve nor worsen.

The weather conditions during the drives differed to a large extent between participants and between pre- and post tests.

The quality of the assessment

When assessing the quality of driving (safety, suppleness, etc) the "expert eye" is superior to any so-called objective measurement using instrumented cars or driving simulator performance. In this study, "real life driving" has a high face validity in comparison to other methods. For this reason, the feedback drive and the expert assessment are strong features of this study. Despite this, there are three elements in this study that are a threat to the validity of the assessment:

On the pre-test feedback drive of the experimental group, the expert assessors were aware of the fact this was the pre-test drive of a person of the experimental group. Knowing that this group was in need of training, this knowledge may have influenced their judgment and expectation, leading to a rather negative assessment.

On the post-test feedback drive, the assessors were aware of the fact that these were post-test feedback rides. A strong point was that they did not know whether a participant belonged to the experimental or to the control group.

The pre- and post feedback drives were not assessed by the same instructor. It is unclear to what extent this may have led to differences between groups.

The research design: the task of motivating the control group

A problem in using a control group, is how to keep the control group interested in the study. A demotivated control group is not a good control group. In this study it was unavoidable to frustrate the control group, as all participants signed up for the "challenging safety course".

In order not to lose too many members from the control group at the outset of the study, we chose not to communicate to the control group that they were the control group. Instead, we spoke about participating in different courses. Although it led to confusion and demotivation on the training day itself, we expected that refusal rates would have been much higher had we mentioned the control group membership earlier in the study.

The question remains whether the control group was too frustrated with their “training” to allow for a valid measurement. The satisfaction questionnaire, however, shows that 83% of the control group agreed or highly agreed with the statements that the pre-training feedback drive was "useful", and that 90% agreed or highly agreed with the statement that the drive was "fun". In this respect, the control group's responses did not differ from the experimental group. Compared to the pre-training feedback scores, the scores of the control group on the post-training feedback driver were even higher. This leads us to conclude that the control group was highly motivated and, in this respect, did not significantly differ from the experimental group.

4.2 The results of the study

The results are primarily based on the effect study.

How attractive is the training for those who did not participate?

Of 500 potential participants, 28% did not sign up for the course, 40% did not show up for the pre-test feedback drives, and 7% did not participate in the post-test feedback drive. About 25% finished all elements. From these figures it might be concluded that the majority of young drivers (68%) are not interested in such a course, even when a) all travel expenses are paid for, b) the course is free of charge, and c) incentives are used (lottery). The motives young drivers

give for not signing up (no time, don't feel like it, not interested), also present indications in this direction.

How attractive is the course for those who did participate?
All the participants that completed the course filled out a questionnaire asking how useful and fun the different components had been. The results show that both control group as well as experimental group are similar on their judgment of the feedback drive. The experimental group also had a track training and a group discussion. The latter gets the lowest rating, while the track training gets the highest. However, the relative minor difference with the feedback drive opens up the discussion about the necessity of track training in these safety programmes. It is frequently mentioned that track training is needed in order to provide an attractive programme. The findings in this study lead to the conclusion that attractiveness is not dependent on the presence of track training after all.

Although exactly the same training programme was offered on two locations, the satisfaction scores differ between locations, both on the track-training programme and in the group discussion. The information from the process evaluation demonstrated that the trainers on the Lelystad track site were not happy with the assumptions underlying the programme. For instance, the programme emphasized that manoeuvring skills should not be trained, but that risks should be demonstrated in order to increase self-awareness about risks and to avoid over-confidence. The ANWB trainers, however, were convinced about the value of skill training. From participants we learned that these views were expressed during the course, possibly leading to confusion and dissatisfaction in the participants. This finding stresses again the importance of the personal views of the trainer for the success of the education programme, as has already been demonstrated in other educational fields, and has been documented by the ADVANCED project.

The power: consequences of a smaller sample
As only 33% of the participants actually completed the programme, the power of the study in order to be able to detect changes (given one is there) has been greatly reduced. From the power analysis it was concluded that the comparison between "regularly" trained candidates and the candidates that had followed a "best practice education" had to be dropped from the analyses because of too few participants in each cell.

Has the training resulted in a change in knowledge, attitudes, and behaviour with respect to the objectives of the course to increase "risk awareness"?

Effects on Self assessment, risk awareness and calibration
The questionnaire measured the effect on risk awareness and self-assessment of skills. No statistical differences were found on the post-test between the control and the experimental group. To rule out the possibility that the questionnaire was not sensitive enough to find differences, the data were also analysed for gender differences. In line with expectations, differences showed up between men and women. Men were more confident and positive about their driving styles and saw less hazards. This, however, did not change after the training. This led to the assumption that the questionnaire measures (more stable) personality traits and attitudes which can not be altered by a one-day training course.

There were some significant effects in the situation judgements, but these were somewhat contradictory. There was only an effect in one out of two situations. However, the control group, which should theoretically remain at the same level, performed worse after the training. Moreover, there is no explanation why the group without any intervention improved in the same way the experimental group did.

Does the training have an effect on those young drivers who are most at risk?
The training's objective was to improve calibration. Calibration is defined as being the balance between self-assessment of skill and risk awareness. Those youngsters that are high on self-assessment (who think that they are extremely good drivers) and have a low risk awareness (regard dangerous behaviours as not dangerous at all) are particularly at risk. Another group that needs particular attention is the group highly insecure drivers with a high risk awareness. The training did not significantly affect either of these two risk groups, with respect to their self-perception and risk awareness.

Effects on self-assessment (on-road observation form)
The participants were asked after the on-road feedback drive about the quality of their own driving with respect to vehicle control, safe and defensive driving, and anticipation. The training did not have an effect on the assessment of skill or complexity of the driving task, according to the young drivers themselves. This is, on the one hand, positive because the participants did not think they improved their skills as a result of the training. On the other hand, there was no change in the perception of the complexity of the driving task.

It is possible that the young drivers already had a very accurate image of their quality as (car) drivers and about the complexity of the traffic situation. To study this the scores of the participant were compared to the scores the instructor had awarded them on the same 4 items;

vehicle control and general skills

safe and defensive driving

anticipation

estimation of the complexity of the driving task

On "vehicle control and general skills", instructors and participants did not differ in their assessment neither on the pre-test nor on the post-test. On "safe and defensive driving", in the pre-test, participants rated their performance higher than the instructor did. As the course was directed at improving self-assessment skills, one would have expected this to improve in the sense that their assessment would be more in line with that of the instructor. This was not the case.

Driving assessment
The driving performance of the participants was assessed by the driving instructor during the on-road feedback drive, before and after the training. Three aspects of driving were assessed:

vehicle control

driving skill

calibration

The results on the driving assessment show that there was a significant effect of the training on driving skill and calibration skill. This is visible in the difference between experimental and control group, but with some analytical complications and inexplicable patterns. The biggest problem is the difference between the experimental and control group during the first feedback drive. The positive effect of the experimental group and negative effect of the control group can be explained completely by the absence of this difference during the second feedback drive. The difference during the first feedback drive is probably caused by the fact that the participants in the experimental group had to drive in a unfamiliar environment, whereas the control group could drive nearby their homes; most of the times in the same place were they got their drivers licence. Although this undesirable phenomenon can be explained, it still interferes with the interpretation of the effects of the training.

There also seems to be a difference between the performance in the two locations, Lelystad and Rijssen. The performance of the participants on calibration skill increased after the training in Rijssen, but decreased after the training in Lelystad. This time the performance of the participants during the first feedback drive is exactly the same for both locations. Therefore the difference between Lelystad and Rijssen can be attributed completely to the effects of the training.

The process evaluation indicated that the trainers in Lelystad and Rijssen did not share the same opinion on the definition of a useful training. Without drawing any conclusions as to who is right in this matter, it is very possible that the trainers in Lelystad had to give a training which they basically did not agree with. This could have (subconsciously) affected the way they gave the training, or the way the participants perceived the training. Research has shown (ADVANCED, 2002) that any education, looses its strength if the educator is not absolutely convinced about what he/she is teaching. Moreover, that the effectiveness of the education is largely dependent on the person, the beliefs of the teacher, and his behaviour (Hale and Glendon, 1987). For a more detailed discussion of the role of the "teacher", see the ADVANCED report.

5. Conclusions

In the Dutch pilot, the recommendations of the ADVANCED report were closely followed with respect to the content of the course and the evaluation of its effects. However, in practice these recommendations were not always followed, as was the case in one of the two locations.

In this study, it has been demonstrated that, on the one hand, the second phase is recognized by the participants as a useful and necessary part of their driving career. On the other hand, the high refusal rate demonstrates that youngsters are not interested in participating on a voluntary basis. The effects of the course are limited, and can even be negative, if trainers are not fully equipped to give the course, indicating that a much greater effort is needed in training second phase trainers than has been the case in this project.

In modules that were delivered by extra qualified trainers, effects were found in the self-awareness calibration factor of the on-road feedback drive. Moreover, participants also valued these modules highly. On many other measures -including the questionnaire- the course did not result in any changes, which in itself leads to the question on how to interpret the results. Are the effects found real effects, and are the other instruments insensitive to measuring these effects? Or alternatively, are the statistically significant effects not to be found in real-life traffic? Answering this question is only possible when more evaluation studies of this type of training are carried out.

References

The EU ADVANCED Project: Description and Analysis of Post-licence Driver and Rider Training. (2002). Rijswijk: Commission Internationale des Examens de Conduite Automobile CIECA.

Brown, I. D. (1989). How Can We Train Safe Driving? Haren: Traffic Research Centre, University of Groningen.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. Hillsdale NJ: Lawrence Erlbaum Associates.

Deery, H. A. (1999). Hazard and Risk Perception among Young Novice Drivers. Journal of Safety Research, 30(4), 225-236.

Engström, I., Gregersen, N. P., Hernetkoski, K., Keskinen, E., & Nyberg, A. (2003). Young Novice Drivers, Driver Education and Training. (Report No. 491A). Linköping, Sweden: Swedish National Road and Transport Research Institute.

Hale, A.R. & Glendon, A.I. (1987) Individual Behaviour in the Control of Danger Industrial Safety series, vol 2. Elsevier, Amsterdam

Hatakka, M. (1998). Novice Drivers' Risk- and Self-evaluations. Doctoral dissertation, University of Turku.

Kuiken, M. J., & Twisk, D. A. M. (2001). Safe Driving and the Training of Calibration. (Report No. R-2001-29). Leidschendam, The Netherlands: SWOV.

McKenna, F. P., Stanier, R. A., & Lewis, C. (1991). Factors Underlying Illusory Self-Assessment of Driving Skill in Males and Females. Accident Analysis and Prevention, 23(1), 45-52.

Stevens, J. (1996). Applied multivariate statistics for the social sciences (Third ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Willems, B., & Cuyvers, R. (2004). Ervaring en Ongevalbetrokkenheid. (Report No. RA-2004-30). Diepenbeek: Steunpunt Verkeersveiligheid bij Stijgende Mobiliteit.

Appendix 1 Organizations involved

- Gelderland Provincial Road Safety Board (project leader).
- ANWB (training Lelystad).
- BOVAG (involved in the project with regard to the role of driving schools in the training programme)
- Central Licensing Bureau, Eastern region (organization of the feedback drives).
- The Ministry of Transport (involved because of possible introduction of compulsory second phase training in the Netherlands)
- FAM (involved in the project with regard to the role of driving schools in the training programme).
- Novem (involved because of the integrated part of the "New Driving" training programme)
- SWOV (realisation of the evaluation research).
- Traffic Test (organization and guidance of the training programme).
- VVCR (training Rijssen).

- SPAIN RACC

NovEV
Results of the RACC second phase

Albert Aluma, Lluis Puerto, Jaime Sanmartin, Gabriel Molina

RACC MANAGEMENT	**INTRAS[(1)] MANAGEMENT**
Project manager: Mr. Pere Sauret **Training design & track co-ordinator:** Mr. Albert Alumà **Project & Training organization:** Mr. Lluís Puerto **On-road co-ordinator:** Ms. Miriam Monfort	**Evaluation team:** PhD. Jaime Sanmartín PhD. J. Gabriel Molina **Training consulting:** PhD. Francisco Alonso **Workshop trainer:** PhD. Francisco Toledo (1) Institute of Traffic and Road Safety. University of Valencia. intras@uv.es http://webintras.uv.es

1.1 Summary of project

The NovEV pilot trial of the RACC Automobile Club took place during the period between January 2003 and May 2004, and involved 621 participants from three provinces of Spain: Madrid, Valencia and Barcelona. The aim of the project was to evaluate a post-license training course in order to assess if it can positively influence the behaviour of novice drivers. An evaluation strategy was planned, based on a experimental research design with experimental and control groups assessed at two points, before and after the training, with regard to a number of variables related to safe driving.

The partner structure that RACC built in order to develop and implement the NovEV pilot trial was led by the European supervisor, CIECA, whereas at a national level, the traffic authority DGT (Dirección General de Tráfico) supported the pilot trial and will use the results for the future development of post-license courses. The training design and its implementation at national level was managed by the RACC Automobile Club and INTRAS (University of Valencia).

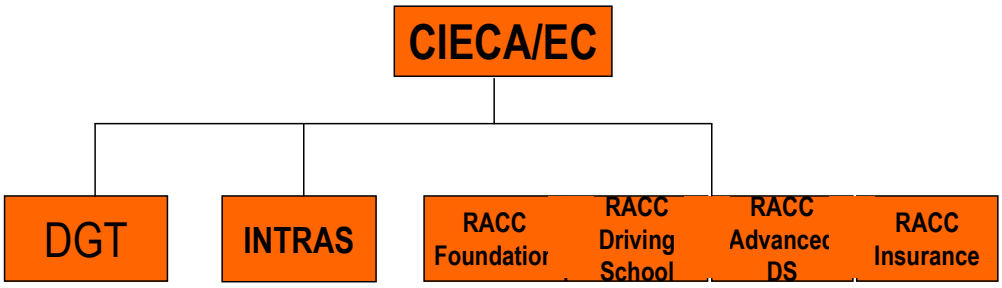


Figure 1: NovEV pilot trial: structure of partners

In January 2003 a massive marketing campaign by post was addressed to more than a thousand young drivers who were all policyholders with HDI car insurance company. These potential participants fulfilled the following specifications: aged 18 to 24, less than 3 years driving experience and living in Valencia, Barcelona or Madrid provinces. The letter informed them about a pilot trial in which they were invited to participate and explained what it involved (at least two tests over a one year period) and what would they get (possibility to take a training course, and to win a car in a lottery).

Phone calls followed the marketing campaign in order to recruit participants and to conduct a short interview-questionnaire (**see 1.7, selection questionnaire**) that would provide the basic background from each participant. The information taken from the phone questionnaire supplemented the information from the insurance company database which provided the basic variables needed to segment the sample into two balanced groups in terms of age, gender, educational background, driving experience and vehicle use.

At this stage two groups were formed: a control group and a test group that totalled 621 people. A pre-test (**see 1.7; pre-test**) of driving behaviour was sent to them and 350 answered within the deadline (183 from the test group and 167 from the control group). Due to an unexpected high rate of dropouts, the participation in the project was extended not only to HDI insurance holders but also to other members of the public, mainly recruited at driving schools in Valencia and Madrid.

The training days took place in Barcelona (3 days), Valencia (3 days) and Madrid (1 day) in July 2003, involving the 183 members of the test group. It consisted of a one day training during

which all participants had to take three areas of training: track training, on-road feedback drive and psychological workshop. Before the training began, the RACC conducted a rehearsal day in order to train the trainers and to improve certain organisational aspects.

The track part consisted of two parts: performing emergency braking with and without ABS on slippery and rough surfaces on one hand, and experiencing an exercise in which participants were distracted by mobile phones and peer pressure. The on-road section combined urban and rural roads on a pre-defined circuit in which each participant had to drive for 20 minutes. In the workshop section, the most important sociological and psychological aspects that affect young drivers were introduced and discussed. Every session except the workshop had a feedback session during which the trainees had the opportunity to interact with the trainers and to have their questions answered.

According to the project guidelines, a period of 5 months then elapsed during which no contact was had between RACC and the participants, in order to allow for consolidation of any attitudinal improvements as a result of the course among the members of the test group.

During the period from December 2003 to end of January 2004 the post-test (**see 1.7; post-test**) was sent to the 350 people still involved in the NovEV pilot trial. The final participation figures after dropouts from the pre-test and the post-test was 263 novice drivers, namely 126 from the test group and 137 from the control group, of which 66% were from the HDI insurance database and 33% were from driving schools.

During the period from January 2003 to March 2004, any reported accidents were monitored amongst participants from the HDI insurance database. Despite being an unstatistically consistent result, the query showed that 4 participants from the control group were responsible for an accident, whereas only 1 participant from the test group was responsible for an accident.

Two basic methods have been used to collect the data from participants: phone interviews (recruiting questionnaire) and post (pretest and post-test). One last source of information has been the database from HDI.

After all the data was collected, a comprehensive statistical analysis was carried out on the data from the final sample of participants that completed the two driving behaviour tests. The original data provided by the HDI database and the recruiting questionnaire were also used to perform the analysis by providing segmentation variables and to detect any self-selection bias.

The evaluation strategy was based on an experimental research design with experimental and control groups assessed at two points, before and after the training, according to a number of variables related to safe driving. An univariate ANOVA model was used to analyse the data of our mixed between-within design for each one of the five scales considered.

Data analysis results showed statistically significant differences between the control and test groups for the “Skills for Careful Driving” scale, meaning that the mean score in this scale was higher for the test group than for the control group after the training. This result goes in the expected direction given that, as reported in earlier studies, self-evaluation of skills for careful driving is inversely related to accidents. Positive differences between the test and control groups were also found for the other four driving behaviour scales, but these differences did not appear to be statistically significant, so they could have occurred by chance.

Finally, data analysis of the course feedback obtained from the participants of the test group showed a rather positive evaluation of the course and the course results. The first conclusion is supported by mean scores over 4 (in a 1 to 5 scale) for the items related to the course organisation, contents and tuition; the latter through the mean scores for the improvements

which were reported by participants, which were significantly higher for the items related to self-awareness about risks and bad driving habits than for driving techniques and skills.

1.2 Participants

In order to have a wide representative sample of Spanish novice drivers, the RACC's NovEV pilot trial was developed by recruiting participants from three major provinces of Spain: Madrid, Barcelona and Valencia. These provinces include both rural areas and urban metropolitan areas. The overall objective was to conduct a pilot trial with 512 participants (256 for each group; see Table 1), divided by 6 quotes of age and gender for both the control and test groups. In order to equal sized groups at the end of the experiment, the control group was enlarged to compensate for expected dropouts.

Table 1: Sampling design with the defined target quotes.

TARGET					
Gender	Age	Barcelona	Madrid	Valencia	TOTAL
Female	19	20	7	20	47
	20	20	7	20	47
	21	20	7	20	47
	22	20	7	20	47
	23	20	7	20	47
	24	20	7	20	47
		120	42	120	**282**
Male	19	20	7	20	47
	20	20	7	20	47
	21	20	7	20	47
	22	20	7	20	47
	23	20	7	20	47
	24	20	7	20	47
		120	42	120	282
		240	**84**	**240**	**564**

The main source of participants came from the database of HDI, a motor insurance company that operates in all three provinces. An initial query into the database provided a rough sample of 5593 potential candidates that fulfilled the three required conditions:

- Geographical location: Madrid, Barcelona, Valencia
- Driving experience: maximum 3 years holding a driving license
- Aged between 18 and 24 years old

As Table 2 shows, the database population was not uniformly distributed with regard to the quotas required. The problem was mainly with younger women from Madrid and Valencia, therefore, some of these quotas had to be filled with subjects from the closest age categories.

Table 2: Sample of candidates provided by HDI database

POPULATION	Age	Barcelona	Madrid	Valencia	
Gender	19	94	3	14	111
Female	20	160	4	17	181
	21	237	8	20	265
	22	362	21	34	417
	23	410	17	52	479
	24	388	20	36	444
		1651	73	173	1897
Male	19	403	8	73	484
	20	699	43	95	837
	21	804	30	86	920
	22	661	23	50	734
	23	370	12	34	416
	24	278	8	19	305
		3215	124	357	3696
		4866	197	530	5593

The marketing campaign by post sent letters to approximately 1300 candidates and asked them to participate in the pilot trial. In exchange, some of the participants that took part would receive a free training day in advanced driving (50% of them in total) and also would participate in a lottery, in which they could win a car. Following this letter campaign at the end of January 2003, the first phone contact with potential candidates came around 1 month later in order to recruit those who had previously been informed by the letter.

The phone campaign was conducted by an external service provider that had been previously informed of the two objectives of this telephone questionnaire:

- Establish if the candidate wished to participate.
- If so, conduct a telephone interview using a selection questionnaire.

All participants were told that they would have to respond to two driving behaviour tests (pre-test and post-test) that would be sent by post or e-mail, with 8 months in between. In the meantime, there would be the chance for 50% of them to take a training course for free. In addition, they were informed that, after completion of both the pre-test and post-test, they would automatically be included in the car raffle. After this "letter+phone call campaign", some gaps left to satisfy our planned sampling design. Thus, in order to get the sample size we had hoped for, some additional candidates were taken from the databases of driving schools in Valencia and Madrid, namely the areas where the HDI database was having the least success. Finally, an initial sample of 621 candidates accepted to participate in our research.

To ensure an appropriate balance between test and control groups when splitting the total sample into both groups, some relevant variables were considered: on the one hand, the selection questionnaire included questions such as educational background (3 levels) and driving experience (mileage per year); on the other hand, other variables like gender, age, licensing date, etc. had been acquired through the HDI database.

The initial test of driving behaviour, or pre-test, was sent out in April 2003 by post or e-mail to the 621 participants that had agreed to participate in the research. By mid-May a reminder letter was sent to the vacant responses. Some of the participants even answered the pre-test immediately before starting the training course. No responses were admitted after the training courses of July 2003. In total, the response rate for the pre-test was 66%, namely a total of 350 participants.

Experience after the pilot trial showed that the dropout rate is quite high and that it is particularly high at the beginning of the project, as the figures indicate:

Recruiting rate: 621/1300: 48%.
Pre-test response rate: 350/621: 56%
Post-test response rate: 263/350: 75%
The post-test response rate is considerably higher taking into account the period of time between pre-test and post-test during which there was no contact between the RACC and the participants.
Pre-test 1st sending: April 2003
Pre-test reminder: May 2003
Pre-test collection: May to July 2003
Post-test 1st sending: December 2003
Post-test reminder: January 2004
Post-test collection: January to March 2004

1.3 Training Programme

A group of at least 36 people per training day was required for the training course. 36 people was considered by RACC to be the minimum number necessary in order to state a positive business case for the future nation-wide implementation of the training.

As people arrived at the training site, they were first received in a welcome session, during which the participants were informed about the organisation of the training, without creating any expectations about what they would experience. Afterwards they were split up into 3 groups so that each group would be participating in one of the three programmed activities (feedback drive, track training or psychological discussion) at any given time. The three activities took place on a rotational basis, so that all sessions took place simultaneously, which was useful from an efficiency perspective. In the working day timetable below, the different colours represent different exercises (yellow-road exercise, blue-track exercise, red- psychological workshop and green-welcome, lunch and conclusions) and the capital letters represent the 3 groups that are on a rotational basis.

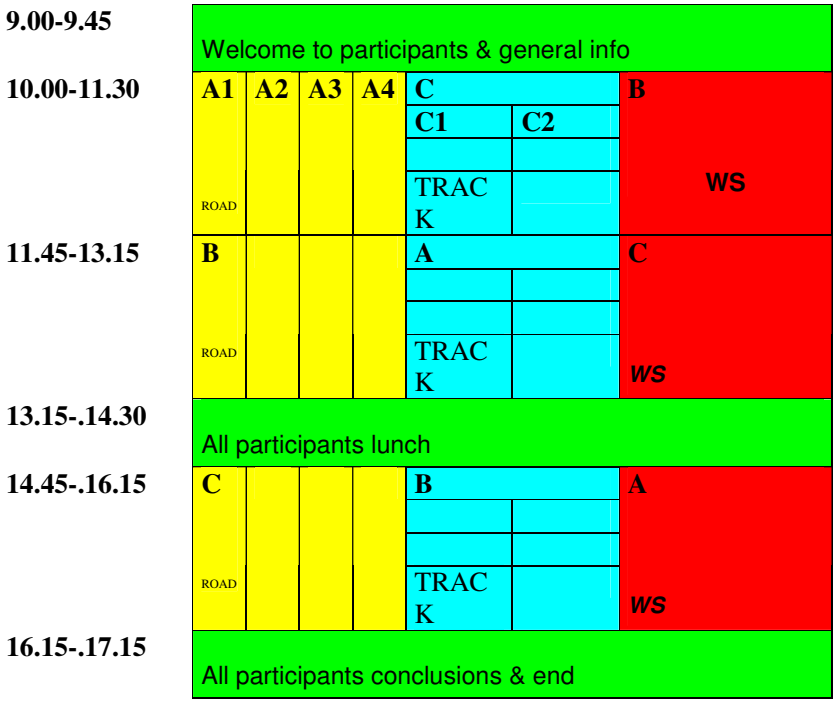


Figure 2: Timetable of the RACC training course

1.3.1 Description of the training

1.3.1.1 Track session

The track session was split into two main exercises, an emergency braking exercise and a slalom (with some distractions inputs). These exercises took place at the same time, so the group of 12 participants taking part was divided into two subgroups of 6 people each. The activity lasts 90 minutes so each group of 6 people stayed 45 minutes in one activity and the next 45 minutes in the other one.

1.3.1.1.1 Emergency braking exercise

This exercise is now explained, as follows:

Goals of the exercise
-To raise awareness of risks related to movement (speed) and stopping (braking).

Operational goals of the exercise
-To teach the correct way to brake
-To raise awareness of the difference between a normal and an emergency braking
-To make participants aware of their individual capacity to brake
- To raise awareness of how speed influences the control of the vehicle direction and coming to a halt.
- To raise awareness of the different road surfaces (especially related to weather conditions)
- To raise awareness of how the driver's physical and psychological condition affect the braking distance due to the reaction time.
- To raise awareness of the limitations of active safety systems helping us to control the vehicle and braking.
- To raise awareness of the limits of the vehicle (tyres, shock absorbers, brakes, etc)

Risk Factors
-Safety distance (time reaction and breaking distance)
-Speed
-Environmental factors (specially weather conditions)
-Psychophysical conditions (transitory psychophysical disorders)
-Over-reliance on active safety systems
-Conditions of the vehicle (tyres, shock absorbers, brakes…)

Messages
-Stopping a car is more important than accelerating
-The safety margin is the main guarantee for a correct braking process
-Emergency avoidance manoeuvres can be dangerous
-Speed is a risk factor that influences the braking distance
-The characteristics and conditions of the road influences our braking distance
-Our reaction time is not the same everyday.
-In-car technological systems can help us but they cannot work miracles

Counterproductive conclusions and messages
-Our skills can compensate for the risk factors
-By means of practice we can compensate for the negative effects of speed or the risk factors of the road
-The active safety systems of the vehicle can compensate for the risk factors
-"Because of my strong avoidance skills I can't crash"

-"Since I did it and nothing happened, never will ever happen"

Description of the exercise
At the start of the exercise, in groups of six people, the participants are given a field form (see 7.7) to be completed by them with their personal details responding to four questions featuring on the first part of the paper.

Participants are free to exchange ideas and opinions with each other.

A – What is the difference between a normal and an emergency brake?
B – What are the differences in the actions that the driver has to do (feet and hand movements)?
C – Can ABS help? Discuss if it is positive or negative

Once the first three questions have been answered, they have to mark in the picture of the field form (see 7.7) where the participant estimates that the car will stop, considering that the driver starts to brake at 40 km/h at a specific given point.

Between the two surfaces (smooth and rough) we will have some foam cubes with letters that indicate several braking distances. The trainer will know the distance between the braking line and the cube.

The participants will have to estimate three braking distances:

- With ABS on a wet rough surface
- With ABS on a wet smooth (slippery) surface
- Without ABS on a wet smooth (slippery) surface

All the braking exercises were planned to be on a wet surface in order to avoid differences in the exercise conditions if, for instance, it rained during one of the training courses and not in others, to avoid tyre screeching, to avoid the feeling of being in a competition, as well as to prevent damage to the vehicle tyres. Only the distraction exercise was performed in dry asphalt conditions since the aim of the exercise and its conclusions could not be spoiled even if it rained.

Everyone drives the car and brakes, observing where the car stops. When the participant stops after braking, the trainer comes up and they compare the real stopping point with the point previously indicated by the participant on the form, evaluating the distance in metres and the error margin.

It is important for participants to evaluate the speed reached in each exercise and we insist on them braking on the braking line.
To avoid wasting time getting in and out of the car, every participant will try the three exercises one after the other, leaving the slippery one until the end.

Feedback
Once the group has finished the exercises and has experienced it for themselves, there is a discussion focusing on how difficult it is to assess braking distances and other things we do not normally consider when driving.

If the concept of normal braking has not been discussed or asked yet, it is time now to introduce the concept. Normal braking is when the speed, the available space and foresight do not require a specific technique: only braking softly in the available space and engaging the clutch at the end is enough.

Questions to the participants in order to start the discussion include:
- Do you think you guessed (the braking distances) right?
- What kind of dangers do these situations have?
- Ask the participants how many correct answers they had.

Closing of the exercise
A safe driver is able to anticipate dangers, does not underestimate risk and does not over-rely on having total control of the vehicle.

Closing message: Driving ignorance can kill you
Life = Education (education is life)

Additional issues derived from the loss of trajectory in slippery surface braking

Goals of the exercise
To raise awareness of risks connected with loss of adherence and skidding.

Risk factors
-Unexpected external factors like different road surfaces cause loss of trajectory

Messages
Speed is a risk factor and can have a major impact on surface adherence

The conditions and characteristics of the road affect our capacity to maintain the stability of the car

Counterproductive conclusions and messages

-No one can know the conditions and characteristics of the road unless you usually drive on it.
-Countersteering corrects the skidding. Skidding is no problem as long as I can countersteer.

Description of the exercise
The smooth surface is 3,5 meters wide. Sometimes the car can lose its trajectory when someone brakes without ABS and blocks the wheels. It will cause a wheel to run into the rough surface – with more adherence – thus making the car spin.

Feedback
Once everybody has been surprised by this situation, it is important to be attentive to their comments. The trainer must stress that a lack of previous driving training will produce hesitation and confusion in the driver when such a situation occurs in reality, and that he/she will not be able to handle the situation successfully. The aim is to educate trainees to identify and prevent problems before they occur by making them "live" the situation, and not to solve it once the problem occurs.

Self evaluation of the participant and interaction with the trainer
The participant will only have one opportunity to experience this exercise and the trainer will evaluate what the participant has actually done.

The trainer asks the participant about his/her feelings during the exercise.

If the participants insist on "seeing" the "correct "solution, the trainer must explain to them that the course is designed to highlight risk situations, and not to teach participants how to live with a high probability of an accident.

Closing of the exercise
A safe driver is able to anticipate dangers, does not underestimate risk and does not over-rely on having total control of the vehicle.

Closing message : driving ignorance can kill you
Life = Education (Education is life)

1.3.1.1.2 Slalom (distractions)

This exercise is now explained, as follows:

Goals of the exercise
Rising awareness of the risks connected to speed/space and the importance of concentrating (avoiding distractions) while driving.

Operational goals of the exercise
-To raise aware of the correct body positioning behind the steering wheel (including hands and feet).
-To encourage self-evaluation.
-To raise awareness of speed limits your options.
-To stress the importance of "sustained attention" when driving and the problems that lapses in concentration can cause.

Risk Factors
-Inadequate position of hands and body
-Road (its limitations)
-Speed
-Lapses in concentration (e.g. mobile phone) and others
-Physical and psychological conditions

Messages
-The correct position of hands and body is very important for safety reasons
-Concentration is an important aspect of safe driving
-Our ability to concentrate is limited
-Distractions while driving affect our attention and increases the possibility of an accident.
-Mobile phone are one source of distraction but other distractions can cause the same effects
-"Hands free" mobile phone systems can still affect concentration on driving

Counterproductive conclusions and messages
-Our skills can compensate for all the risk factors
-With practice we can compensate for the negative effects of speed or other risk factors
-The use of "free hands" systems does not cause distraction
-Since I did it and nothing happened, never will ever happen

Description of the exercise
At the start of the exercise, participants are given a questionnaire (see on 7.7) to be completed by them with their personal details and their opinion about mobiles and how other activities affect driving.

The participant will drive the car with one trainer as a passenger. The trainer will pressure him/her to drive as fast as he can (to simulate peer pressure) but to try not to knock down any

cones along the circuit. Simultaneously he must subtract from 100 to 0 in a series of 3 (100, 97, 94, 91, etc)[49] , and this will be monitored by another trainer (outside) through the mobile phone (replaced by a radio station).

The exercise finishes when the participant has completed 6 timed laps. The 6 laps are divided in groups of two laps. The two first laps are to observe the participant, the two next laps will be timed trying to make the participant drive at his/her limit. The two last laps the participant will have to talk with the mobile phone while still driving at his/her limit. The first of the last two laps includes a signpost (a poster held by a trainer which simulates a signpost…) shown to the participant. The subject will have to recall it at the end of the exercise when he/she will be asked about it. The minimum requirement is 12 subtractions per lap with no more than 4 mistakes and at the end an ability to recall the poster.

Feedback
The paper handed in at the beginning is returned to the participants for them to answer some new questions:

-What is your conclusion and perception of the exercise?
-Do you think that you can do more things while you are driving? Justify the answer
-Up to which point could you solve the difficulties?
(They confirm or maintain their opinion according to their level of self-confidence.)
-Before the exercise, did you think that you'd do it better, worse or roughly the same?

Self-evaluation of the participant and interaction of the trainer
The trainers will have to evaluate the participant comparing variables such as the differences in time with and without distractions, the errors made (knocking down a cone or going off the track) and the participant's ability to recall what was on the signpost.

Closing exercise
The driving task (even though it is routine) can sometimes be complicated and requires all of our attention if we want to drive with a high level of safety

Closing message: driving ignorance can kill you
Life = Education (Education is life)

1.3.1.2 On-road session

In the on-road session the participants are going to test their driving ability. The group of 12 people is divided into 4 groups of 3. The groups get into a car with one trainer per car. Then they start the exercise where each participant will have to evaluate his own driving as well as the driving of the other two participants. Simultaneously, the trainer also evaluates the three participants, whenever each individual is driving. Each driver has to follow a pre-determined circuit for 20 minutes, after which there is a general discussion about the strengths and weaknesses of each participant's driving style.

Goals of the exercise
-Think about the risk-increasing factors while driving caused by bad habits, risky driving style, unsafe speed adjustment, vulnerability of other road users (e.g. pedestrians), information overload, weather conditions, insufficient skills, etc.

[49] To evaluate if the participant has problems subtracting, and to distinguish between his mathematical skills and how distracted he is, they are given a paper before doing the exercise on which they are asked to subtract 3 from 100 to 0 in 50 seconds.

-Make a real self-evaluation of one's own driving, including personal strengths and weaknesses, basic driving skills, personal driving style, strong and weak points of their skills when encountering risky situations.

-Become evaluators of the 2 other participants in your car, paying attention to the above points. This can be useful to highlight habits or errors, because peers often have a great deal of influence on each other.

Description of the exercise
The participants know about this module in advance. They will use a driving school vehicle and they will have to fill in a questionnaire.

One driving school instructor with three participants and a driving school car (double pedal system and an oriented rear-view mirror) perform the driving exercise in urban and rural areas, developing their knowledge and skills of all participants relating to:

-traffic rules
-observation and taking into account traffic signs
-anticipation of approaching signs
-speed adjustment
-road positioning
-driving priority
-safety margins

Characteristics of the route
The route is the same for all participants. There are two planned stops to exchange information with the driver and there is a mix between urban and rural roads.

The route must include specific situations connected with certain traffic signals. These situations are: roundabouts / regulated crossroads / non-regulated crossroads / merging into a high speed road / left and right turns / stop signal with visibility / traffic lights / pedestrian crossings. Although the activity will be carried out during the day, participants will be informed about the differences between day and night-time driving (just information).

Description of the exercise
First of all, the instructor explains that the exercise consists of a driving exercise where each participant is going to test his/her own driving ability with regard to its strengths and areas for improvement. The exercise finishes with a self-evaluation.

Each participant has a questionnaire to fill out regarding the driving of the two other participants and the driving surroundings. The questionnaire will be filled in by the 2 observing participants while the other one is driving. The driving participant will have to fill in his/her questionnaire in the 5 minutes following completion of his drive.

Each participant has to drive during 20 minutes, while the trainer and the rest of participants are filling in the questionnaire related to the driver and the surrounding area where they are driving (the trainer and the participants questionnaire are different). See section 7.7.

Secondly, once the 60 minutes have finished, it is time to start a group discussion lasting 30 minutes, and divided into two parts for each participant:

-In the first part, each participant gives his opinion on of his/her own driving and the other participants discuss these opinions. This should be done for all 3 participants.
-In the second one the trainer gives his own feedback, particularly on issues which were not raised by the participants themselves, concerning knowledge, attitudes or skills.

The trainer should be especially attentive to the perception of risk that each participant has of the wrong driving behaviour and the beliefs associated with this conduct. In this way, the trainer will point out the potentially negative consequences that such conduct can have.

Thus, the trainer will introduce supplementary conditions (transitory psychophysical disorders, behaviour of the other drivers, weather conditions and vehicle faults), that could increase the probability of having an accident. These conditions should then be discussed amongst the participants.

Finally the trainer asks the participants if they have some comments to add, collects the questionnaires, and concludes the session with one last message:

Self-evaluation and assessment of the behaviour of other road users is a good thing, thus realising that experience is no guarantee of safety, and that the (health and economical) consequences of accidents for oneself and for others are very high.

A safe driver is the one who knows that incorrect behaviour can sometimes cause an accident.

Closing message: driving ignorance can kill you
Life = Education (education is life)

1.3.1.3 Workshop

In the workshop, the group of 12 participants is not divided as in the other sessions. It takes place in the classroom (an itinerant truck in the case of the RACC training) where the workshop trainer does his job. The trainer has to try to keep the attention of all participants during the full 90 minutes of the session

Contents of the workshop
The contents of the course are structured in five blocks with the following titles:

1. Traffic accidents: Statistical information, general types of accident per reference group, what is driving about?, major causes, consequences of accidents.

2. Mistakes and offences: Two facts that cause accidents (strategies to fight against group pressure which can lead to traffic offences)

3. Risk factors in traffic accidents: (alcohol and drugs / drowsiness and tiredness)

4. Use of passive safety systems: reducing the consequences of accidents.

5. Closing and commitment

Developing of the contents of the workshop (goals):
The goals of the course, based on the above content, are as follows:
a)To understand the human factor in the causes of accidents
b)To introduce the concept of incident before the accident. It will be treated us to change the perception of the risk incident and modulate the reinterpretation that we use to do of these ones.
c)To analyse how to avoid accidents by eliminating chance as much as possible.
d)To raise awareness of the consequences (disabilities, injuries, deaths) of accidents suffered by accident victims.
Scientific investigations show that the most effective strategy to teach or modify attitudes with a young group of people is not so much by emphasising the fatal consequences of accidents, but rather stressing the consequences of disability as a result of an accident, which victims have to live with for the rest of their lives.
e) Positive consequences that are less likely to happen than people think

f) To considerate traffic offences as a result of voluntary conduct.
g) Participants have to recognise their own erroneous behaviour in driving offences.
h) False beliefs about risk conduct treated in the programme.
i) Norms and values.
j) Risk perception. To develop the ability to recognise potential dangers in advance, in order to avoid the subsequent risky situations.
k) Overestimation of skills.
l)Peer pressure. In this way the passenger can have a negative influence on the behaviour of the driver.
m) Inadequate ability to fight against social pressure.
n) To make the subject aware that he can control the traffic offences that cause accidents.
o) To increase self-esteem and the self-efficiency of the participants.
p) The feeling of self-efficiency, feeling of competition and capacity produced when the subject can deal with different situations correctly.

Specifically for the closing session
The trainer presents the following scene: Imagine that you have had an accident and it causes you a disability

"Do you think that is justified that by committing an offence you have to suffer the consequences (namely of being handicapped) for the rest of your life?"

There follows private and public discussions on this issue and answers from the participants

Then participants are asked:

"What level of commitment are you going to assume in your future driving career?"

1.3.1.4 Sequencing of the contents

The contents of the programme, according to the five main themes, take place according to the following timetable:

Table 3: Workshop session contents

	TITLE	TIME
1	Traffic accidents: Statistic information, general type of accident of the reference group, what is driving about?, major causes, consequences	15 minutes
2	Mistakes and offences: Two factors that causes accidents	15 minutes
3	Causes of traffic accidents: Risk factors (alcohol and drugs / drowsiness and tiredness)	40 minutes
4	The only way to reduce the consequences is to use passive safety systems	10 minutes
5	Closing and commitment	10 minutes

1.3.1.5 Establishment of communicative / education strategies

The workshop has a mixed strategy where the trainer's presentation should be mixed with interaction and the active participation of the group of participants.

Despite having a theoretical character, one part of the presentation moves into group work, with thinking tasks, discussion and brainstorming about the problems treated at each moment.

The aim of this mixed strategy is for the participants to adopt an active role during the session.

The strategy of the programme is formed by facts such as: On one hand, to consider that the behaviour is a direct consequence of our intention to do something, and on the other hand to consider that intention is connected to some psychological variables.

Activation of cognitive strategies
In a lot of situations it is our thoughts, beliefs and motives that determine our intention to act or not to act in a specific way. This means that our intentions can be explained and predicted.

Activation of behaviour strategies
Some forms of behaviour can be initiated by an automatic or non-conscious act (without thought and previous reasoning) and by repetition (experience). The more conduct is repeated, the more it becomes automatic.

This means that some conduct is caused by conditioning or addicted factors. In these cases, the strength of the habit can obstruct the programme effects or the efforts aimed to increase safe conduct.

Activation of emotional strategies
Recent investigations state that when the initial attitude of the participant is not the expected one, the activation of thoughts of a moral nature in certain kinds of situations and consequences is the most effective way to modify risk conduct. In this way, imagining the implications of such situations helps form a strategy to change and incorporate more reasoning and reflection with regard to our vulnerability.

For this reason, the detailed analysis of physical, emotional and sociable causes and consequences of accidents related to people close to them, combined with the thoughts of personally being involved in an accident, form two methods that can modify conducts, thoughts and attitudes.

1.3.1.6 Final closing

In this part of the training, all the groups met in the truck. The goals of this module are to reinforce all the participants have learned up until that point.

The RACC trainer speaks about two main themes:

-He makes a summary of the day, he talks about the exercises that the participants have done and draw conclusions of each exercise.
-He stresses once again the general message of “driving ignorance can kill you”
Life = Education (education is life)

1.3.2 Locations & Facilities required

For the RACC it is very important that the courses be performed in wide open areas, on the outskirts of an urban area.

The places chosen to develop the activities must only fulfil basic conditions to be suitable for a course, since most of the means and materials were provided by the itinerant truck of the RACC advanced driving school. The basic facilities required are:

- 110 x 40 m^2 of asphalt surface is the minimum space required to develop successfully the track session.
- Water supply to perform track exercises.

- Energy supply is not a must in this case because the RACC Advanced Driving School truck can provide its own energy supply.
- The location of the course must be in the periphery of a urban area for the on-road module, since it needs a mixed route containing urban and rural roads.
- The workshop session does not need a specific facility (building, classroom) because the truck acts not only as a means of transport but also as a classroom where the workshop can take place.

In addition, other material is provided by RACC and INTRAS, such as:

From RACC advanced driving school

- Truck (this truck is very important because it allows for the mobility of the course. It is used to bring the material needed and it can be transformed into a classroom to develop the workshop session with the following material: DVD, video, projector, TV, screen...).
- Cars for the track exercises (minimum of 3).
- Extra material for the track exercise like: sprinklers, cones, banners ...

From RACC driving schools:

- Pre-license driving school cars (cars with dual command) for the road exercise (minimum of 4).

From INTRAS

- Educational material to conduct the workshop.

Training Locations

- **Barcelona:** The RACC contacted the city council to develop the project in the open area "Sot del Migdia", in the vicinity of the Olympic Ring. This place had the minimum requirements of space, water supply, and urban and rural road access, so the rest of the materials, energy supply, classroom was provided by the itinerant truck. Days: 4th, 5th, 6th July 2003.

- **Valencia:** Since the training had to take place in the periphery of the town, the chosen place was the Valencia Trade Fair, and the training took place in the parking area of one of its modules. Days: 11th, 12th, 13th July 2003.

- **Madrid:** The heavy daily traffic of Madrid meant it was better to choose a location further away from the capital. Therefore the training course took place in Humanes, 30 Km away from Madrid, but well connected to urban and rural roads. The chosen facility could provide the classroom and a safe environment to perform the track exercises. Days: 19th July 2003

1.4 Selection of trainers

In order to get the best selection of trainers, the RACC took advantage of its structure and experience, from which the track and on-road trainers could be engaged, whereas for the workshop the trainer was provided by INTRAS (Traffic Research Institute of the University of Valencia).

The responsible persons for the organisation and development of the training were provided by the RACC International Studies department.

Lluís Puerto
Núria Albertí

The three working sessions defined in the project depended on different partners:

Track (RACC advanced driving school)
Road (RACC initial driving schools)
Psychological workshop (INTRAS)

Track

The advanced driving school from the RACC selected a team of trainers that had a mixture of experience and youth.

Their experience was essential to ensure that the exercises and the messages were realised and transmitted without shortcomings and that they were communicated easily.

Youth was also considered to be essential to connect with the participants, for them to assimilate the messages and to identify with the trainers.

Only one trainer performed both as track and on-road trainer, but he never performed both tasks in the same training course.

Co-ordinator
Albert Alumà

Trainer team Distraction Exercise
Enric Riera
Xavier Tomàs
Sergi Corominas
Oriol Aragonés

Trainer team Braking Exercise
Juan Carles Mach
Raül Illamola

Road

For the selection of on-road trainers the RACC chose instructors from the RACC driving schools. Such a choice was, firstly, in order to take advantage of the synergies of our club, (pre-license driving school, advanced driving schools), since it covers several training areas. Secondly, we were able to benefit from the local knowledge of RACC employees. This was useful for identifying areas to carry out the exercises and because the preparation of the on-road routes did not require any advance visual checks. Another significant aspect was that it was decided to do the feedback drives with dual command cars, so the co-operation of the pre-license driving schools was essential.

Co-ordinator
Miriam Monfort from RACC headquarters was responsible for co-ordinating all the driving school teachers that took part in the three locations. She was present in all the sessions and she was the responsible for developing the work sessions in the three places.

The team of driving school teachers varied according to the location, but they were the same in all the sessions of at place. Driving schools from Granollers, Valencia and Valladolid were selected for the Barcelona, Valencia and Madrid training, respectively.

The selection of trainers that took part in the project was performed according to their experience, youth and availability in each zone.

For the session in Barcelona, the directors of the RACC driving school of Barcelona and Tarragona, both with more than 20 years of experience in training drivers, were selected along with younger teachers to contribute to the training of trainers for the "SERVEI CATALÀ DEL TRANSIT", the Catalan traffic authority.

In Valencia we had the co-operation of the RACC driving school in this city that offered us the most expert and qualified trainers.

In Madrid a team of trainers came from Valladolid, because it was the most qualified team to develop a successful session because of their experience in driver education and flexibility in adapting to a new training concept.

The trainers in each place were:

Barcelona
Miquel Mora
Lluís Mestre
Ingrid Ballbé
Oriol Aragonés
Maria Aragonés

Valencia
Maria José Rollo
Cándido del Campo
José Luís Moreno
Miguel García Cristóbal

Madrid
José Luís Martos Gómez
Mariano Cortijo Pérez
Víctor Marcos
Augusto García

Psychological Workshop
In the NovEV project the third module was the psychological workshop.

The “Instituto de Tráfico y Seguridad Vial” (INTRAS, University of Valencia) was the responsible for this area throughout the project. They supplied the most experienced trainers in this area. It was very important to obtain the maximum effectiveness in a short session of 90 minutes.

The trainers in this part of the session required considerable persuasive abilities in order to convince and to maintain the attention of the participants.

One important variable that is connected with these skills is that the workshop trainer should be highly motivated and must be totally willing to commit himself to this kind of training.

Another variable which may help to enhance the persuasive skills of the trainers is that he is a similar age to the participants.

Other skills required by the trainers include:

-General knowledge about psychology
-General knowledge about behaviour and driving

-Knowledge about the psychological characteristics of the group for which the programme is addressed

The following INTRAS officials took part in the design of the workshop:

Dr. Jaime Sanmartín: NovEV pilot project manager

Dr. Francisco Alonso: Training description and definition

Dr. Francisco Toledo: responsible for implementing all workshop sessions that took place in the three locations. A very experienced trainer in training dangerous-goods truck drivers and guiding sessions with small groups.

European Supervisors (CIECA)
Although not directly involved in the training, the supervisors from CIECA played an important role in the definition of the course and helped to improve important details with respect to the training feedback sessions. Esko Keskinen, Nick Sanders and Heleen Groot were involved in the EU Advanced project, which gave them a great deal of experience in assessing second phase driver training from a theoretical and practical perspective.

Based on the recommendations from the Advanced project, and the guidelines on training issued at the beginning of the NovEV project, Heleen and Nick made a practical assessment of the final RACC training in Madrid on July 20, 2003. Various improvements were suggested in case the RACC club wished to develop this course for commercial purposes in the future. The suggestions for improvement largely focused on classic areas which have proved to be difficult to implement properly in other courses too. However, the overall structure and implementation of the RACC training showed that considerable thought and reflection had been devoted to its development, and that the most essential Advanced guidelines had been adhered to as far as possible (see 7.5, experiences implementing the training)

Training of trainers
A rehearsal of the training courses and training-the-trainer process was carried out on June 11th and 12th 2003. On the days leading up to the 11th the theoretical background manual for the course was sent to trainers so that they had time to read it carefully before the rehearsal, and so that they could assimilate the concepts and prepare questions for discussion. On the 11th, all the background and exercises to be carried out were explained to them once again, and a range of issues were discussed.

On the 12th June a sample of 12 volunteers took part in a full-day course, during which all of them were trained in the following order: track training, on-road session and psychological workshop. Several shortcomings were identified and discussed, although most of them were of an organisational dimension and only a few about the course design itself.

Track session
During the Advanced project and before the NovEV pilot, track trainers of the advanced driving school of the RACC were informed in several meetings about the theoretical background and progress of the post-license training concept for novice drivers.

The practical application of the NovEV project did not surprise us and the training of trainers focused more on how to complete the exercises in 45 minutes more than what to do during the exercises.
The messages transmitted had a design and description phase supervised by INTRAS officials.

The rehearsal was particularly useful for identifying areas for improvement, such as:

- The necessary space for the braking exercise
- Braking exercise without ABS in dry surface was too noisy (creating a "racing atmosphere" which we were keen to avoid)
- The wet surface exercise is affected by the rain (if it rains), thereby reducing the impact of the learning goal rough/fine versus dry/wet
- How to prevent the vehicles from sliding off the exercise area.
- ABS braking had to be performed before braking without ABS, in order to avoid over-reliance on ABS (the feeling that ABS can compensate for driver errors).

On-road session
In the road training the practical application, like in the track exercise, was specially based on how to the exercise in 45 minutes because there were some external factors that risked making it more difficult. Some points were identified for consideration:

-The duration of this part was identified as a critical factor to follow correctly the timing of the day.
-It was stressed to road trainers that they had to be familiar with the route and its characteristics.
-The road trainers have to be aware of the traffic situation. Depending on the time of the day and the place where the exercise was, there could be unexpected delays in the timing..

Workshop
The workshop rehearsal did not reveal any problems because the session was designed to last the same as the other sessions, and the trainer had considerable teaching experience.

1.5. Experiences implementing the training (feedback from organisers, trainers, trainees)

Apart from the post-training feedback form sent to CIECA by October 2003, a meeting took place between CIECA, RACC and INTRAS after the last training course, in Madrid on the 19th July 2003, during which all the experience gained from the six training days was discussed in common with the observations that the supervisors from CIECA had made during the final training day.

Both CIECA representatives agreed that the training was highly commendable, both in terms of organisation and content, and that it followed Advanced guidelines to a large degree.

The CIECA representatives did, however, give some feedback which could help improve the effectiveness of the training. The feedback is stated below, divided into the 3 main modules of the course: on-road feedback and training, track training and group discussion.

On-road feedback and training

The on-road section is probably not long enough to either assess each individual properly, nor to give the driver enough time to relax. The burden is on the trainer to make sure that the trainees are as relaxed as possible.

If the same route is being used for all the trainees, and more than one car is being used, their departures should be staggered to ensure that the cars are not simply following each other around the route.

If the on-road session is designed to give feedback and training, then the trainer should be giving feedback and advice during the drive itself (not only at the end), including praise for good performance.

Once the drives have been completed, and the trainees and trainer are required to fill out their observation forms, this is an opportunity for real feedback. The observation forms should be used as a basis for discussion, not as an end in itself. Only by discussing and questioning the statements made on the observation forms will the learning experience be reinforced.

RACC may like to consider the following approach to the feedback session at the end of the on-road drive: firstly, the driver him/herself is given the opportunity to comment on his/her own driving. Then the other trainees, in the role of observers, can comment on the driver's performance. Finally, the trainer delivers his own opinions and raises questions and issues in order to make each driver reflect and to reach conclusions, in conjunction with the driver, on how the performance can be improved.

The observing participants may have difficulty filling out such an intricate observation form. From the experience of the NovEV training, the field form of the trainees should be modified, since it demands excessive attention from them when it comes to monitoring the driver, on the one hand, and the surrounding traffic, on the other. In the future, the different issues to be evaluated should be more synthesised into types of driving mistakes, for example.

Track training

In each session, it is important to get the participants involved from the outset. Instead of demonstrating the correct seating position behind the steering wheel, the trainer should ask participants to comment on his seating position, how to improve it and, importantly, why. This gets the participants immediately thinking for themselves.

During the braking exercise, in order to reinforce the conclusion that adequate safety margins are important, it may be useful for the trainer to encourage participants to think about other variables in reality which affect braking distances, such as speed, trajectory, surface, element of surprise, distractions, poor technique, tiredness, etc. Participants should be encouraged to "discover" these elements by themselves.

The feedback session at the end of the braking exercise should be used to find out what the participants felt and concluded from the exercise, not to allow the trainer to tell the participants what the conclusion was. Perhaps the atmosphere would be more relaxed and intimate if the chairs were arranged in a semi-circle around the trainer, rather than in a line in front of him.

As a general rule, the objectives of each exercise should remain a "secret" until the participants discover the objective for themselves. If you tell people what the aim is in advance, you are prescribing something rather than allowing participants to think for themselves and draw their own conclusions.

In the slalom exercise, it would be useful to consider using a participant to demonstrate the correct seating position. Other participants should then comment on how to improve his position, based on the conclusions made earlier when the trainer was sitting there.

Group discussion

The observations on the group discussion were on a general level:

Changes could be made to provide a bit more variety over the 1½ hour period.

Consider the possibility of including a little group task involving, for instance, groups of 3.

The session focused largely on the weaknesses of novice drivers in general. Perhaps it would be a good idea to introduce a way to encourage participants to reflect on their own, individual weaknesses too, using personal experiences to stimulate ideas and discussion.

A 5-10 minute break during the 90 minute session may be a good idea.

1.6 Evaluation design and timetable

A one-year project was considered necessary in order not only to implement the training course, but also for the pre-test and post-test posting and collection, as well as to allow for a minimum time period before and after the training to let the participants' experiences take affect and thus be able to draw more robust and time-lasting conclusions.

The chart below summarises the process of the NovEV evaluation as designed prior to its implementation.

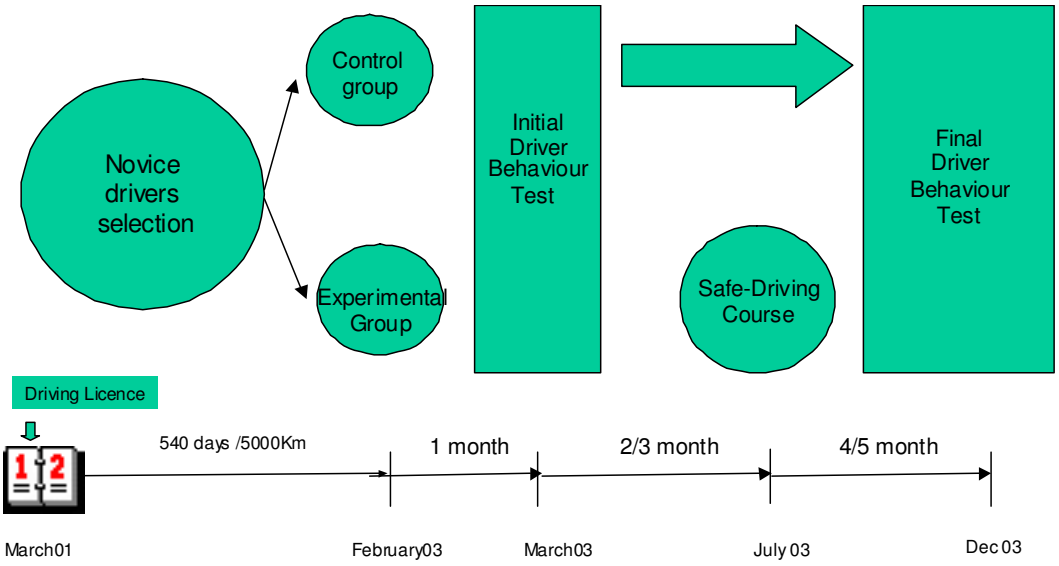


Figure 3: NovEV pilot trial evaluation design

In total, there has been a lapse period of one month between the recruitment of participants and the sending of the pre-test. The collection of pre-test questionnaires was extended from April up to June 2003, including the sending of a reminder in early May, but no more pre-test questionnaires were accepted once the training courses began in July 2003.

Following the training courses in Barcelona, Valencia and Madrid there was a second lapse period during which there was no contact between the RACC and the participants at all. This period lasted for 5 months until December 2004, when the post-test was first sent out, followed by one reminder in January 2004.

After the training and the data collection had been concluded in February 2004, the statistical expert from INTRAS began the data analysis phase. The statistical analysis took place from March to May 2004, taking into account the data collected from the pre-test and the post-test. The preliminary results were presented to the European partners during the meeting in Vienna, in June 2004 and deeper analysis was performed after expert feedback was given by the NovEV independent evaluator, Esko Keskinen.

After this feedback, conclusions were drawn and the final report produced, thus finalising a 17-month project that has been realised by RACC in Spain.

1.7 Data collection methods (questionnaires, on-road feedback form, diaries, registers, etc)

Several types of questionnaires and field forms were used throughout the pilot trial to collect data from the participants. All of them have been used in the statistical analysis, except the field forms of the training programme that were collected as consultative documents only.

Motor insurance company database

The HDI insurance company database provided data about the trainees' vehicle (type of vehicle, value, number of accidents, license date, etc) as well as personal data needed for contacting them (name, phone, address, date of birth, etc). The personal data was only used by the project manager and holder of this data (RACC) with permission from the participants. No third parties had access to personal data.

Selection questionnaire

The first phone contact for recruiting candidates to the project included a short interview that was necessary to build segmentation variables such as: driving mileage per year, educational background (personal and parental), daily driving situations and number of attempts to pass the license examination. There were also basic data questions for those participants that came from the RACC driving schools, rather than from the insurance company database.

Pre-test

The initial test of driving behaviour or pre-test was a multiple choice type questionnaire that all participants had to fill in. It was the main source of data for performing the analysis along with the post-test. The pre-test included driving situations separated in six blocks as follows: driving education, risk factors, strengths and weaknesses, driving situations, driving experience and traffic accidents.

Post-test

The final test of driving behaviour was also a multiple choice questionnaire aimed at collecting data to compare to the pre-test data, in order to evaluate the evolution of driving behaviour of participants both individually and between groups (experiment and control). In addition, it included one section addressed only to those participants that had taken part in the training, in which they were asked to give their opinion on the different parts of the training.

Track training form 1

This field form corresponds to the braking exercise of the track module of the training. It was used not only as support material to conduct the exercise, but also as a feedback form for the participants on which to give their opinion of the exercise. By means of this form, both the driver and the trainer could clearly see the difference between their pre-exercise assessment (estimation of braking distance) and the real braking distance performed.

Track training form 2

This field form relates to the slalom distraction exercise and was used by both the driver and the trainer. The first section was used to test their ability to count down from 100 in series of 3 and to give feedback on the session, and the second section was for the trainer to note the mistakes made by the driver during the exercise.

On-road form 1

By means of this form, the trainees could evaluate not only themselves, but also their two driving partners during the on-road feedback drive. There were three columns with boxes that the trainee had to tick every time the driver made one of the mistakes stated in the form. In this way, the participants had to pay attention to their driving style, as well as to the others'.

On-road form 2

This form was very similar to the first on-road form, but was used only by the trainer. It included one column for each of the three drivers, so that the trainer could have all the information about the trainees' mistakes all together during the feedback session.

1.8 Statistical analysis

The statistical analysis was based on a before-and-after design with control group, which is described in depth in the next section.

2. Evaluation and Results

This section of the RACC/INTRAS report describes the evaluation of the RACC training programme for novice drivers. To accomplish this goal, an evaluation strategy was used based on an experimental research design with experimental and control groups assessed at two points, before and after the training, in a number of variables related to safe driving. Firstly, the pre-test should allow to check that the experimental and control groups are equivalent at the outset of the research; secondly the differences in the post-test for both groups would provide the empirical support to evaluate the impact of the post-licence training on the novice drivers.

METHOD

Participants

Subjects included drivers associated to RACC, an Automobile Club with close to nine hundred thousand members all around Spain. Three conditions were set to be able to participate in this study: age between 18 and 24 years; no more than 3 years driving experience; and residence in Madrid, Barcelona, Valencia, or towns surrounding these 3 large cities. The last condition was set for logistical reasons associated with the organisation of the training. Almost 5800 subjects satisfied the above 3 requirements, which served as the initial database from which to randomly select the candidates who were then given the opportunity to participate in this research.

Finally, 621 subjects agreed to participate in this fist step of the study: 321 men and 300 women (see Table 4 to see their distribution according to age). They were split into control and test groups according to a random assignment blocked by gender in order to keep the number of men and women balanced in both research groups.

Table 4: Gender * Age (reached along 2003) Crosstabulation (N=621)

			Age (reached along 2003)						
			19	20	21	22	23	24	Total
Gender	Men	Count	48	59	55	60	57	42	321
		% of Total	7,7%	9,5%	8,9%	9,7%	9,2%	6,8%	51,7%
	Women	Count	35	41	45	60	59	60	300
		% of Total	5,6%	6,6%	7,2%	9,7%	9,5%	9,7%	48,3%
Total		Count	83	100	100	120	116	102	621
		% of Total	13,4%	16,1%	16,1%	19,3%	18,7%	16,4%	100,0%

Some of the 621 subjects who initially joined dropped out during the different steps planned in our research design. After all, we gained complete records (selection quest. + pre-test + post-test) for only 263 subjects, namely 42.3% of the original subjects who agreed to participate. Table 5 shows the cross-tabulation by gender and age of the final sample of participants.

Table 5: Gender * Age (reached along 2003) Crosstabulation (N=263)

			Age (reached along 2003)						
			19	20	21	22	23	24	Total
Gender	Men	Count	17	22	20	21	22	14	116
		% of Total	6,5%	8,4%	7,6%	8,0%	8,4%	5,3%	44,1%
	Women	Count	19	22	18	28	28	32	147
		% of Total	7,2%	8,4%	6,8%	10,6%	10,6%	12,2%	55,9%
Total		Count	36	44	38	49	50	46	263
		% of Total	13,7%	16,7%	14,4%	18,6%	19,0%	17,5%	100,0%

Measures

We can differentiate between three blocks of measures collected for the evaluation:

(1) Criteria measures. The purpose of the training programme was to encourage safer driving behaviour on the part of the trainees, so accident risk was considered as the general criteria variable to assess in the pre-test and post-test. In order to measure this accident risk, we used a set of scales relating to self-reported driving risk, skills, driving habits and self-evaluation. These scales were devised by Hatakka (1998), who describes the development and application of 15 scales designed to measure (un)safe driver behaviour. We opted for 5 of these scales because of their fine psychometric properties for this kind of evaluation. They were included in the pre-test and post-test questionnaires and given to all the subjects in the control and test groups. Thus, the following five scales were used after being translated into Spanish (note that the short name headings will be used in the rest of the section when referring to them):

- (Scale A1) Self-evaluation of risks connected with personal careless driving habits (7 items; see annex 11 for pre-test/post-test).
- (Scale A2) Self-evaluation of risks connected with showing off and situational reactions (7 items; see annex 11 for pre-test/post-test).
- (Scale B) Self-evaluation of skills for careful driving (7 items; see annexes pre-test/post-test).
- (Scale C1) Self-reported habits related to driving while being in an improper state (7 items; see annex 11 for pre-test/post-test).
- (Scale C2) Self-reported habits related to driving at high speed (8 items; see annex 11 for pre-test/post-test).

(2) Control measures. A set of variables which are usually considered as relevant predictors in traffic research were also collected. These variables enabled us to control the internal validity of the experimental design. They are listed below grouped by affinity:

- About the subject: age; educational level; parents' educational level; years of driving licence; years actually driving; possession of own car.
- About the car: principal user of the car; car power; estimated car value; car age.
- About driving experience: Km. driven per year; number of fines; number of accidents; type of roads most often used (motorway, national roads, urban); main reasons for driving (going out, to work); familiarity with certain driving conditions (bad weather, artificial light, driving alone, etc.); familiarity with driving motorcycles.
- About training interests/motivations: Three scales included in the pre-test indicated the subjects' motivation and reasons for participating in a post-licence driving course, according to 3 main areas: increasing driving knowledge and skills; better able to recognise and avoid risks; and abilities to recognise one's strengths and weaknesses.

(3) Course evaluation measures. Finally, participants in the test group completed a questionnaire, which was included in the post-test (see annex 11 for post-test), designed to get their opinion on some aspects of the training they took: course organisation, trainers' performance, contents interest, didactic resources, usefulness, potential improvements, etc. The same questions were raised separately for the 3 modules of the training course: the on-road feedback and training, the track training, and the discussion group.

Procedure

A near 1300-people random sample of RACC members satisfying the conditions stated in the Participants Section were sent an introduction letter in January 2003. The letter described who we were, our aims and workplan, and included an invitation to participate in the study. Recipients were also informed what would be expected of them if they finally decided to participate, as well as what they would obtain by getting involved: besides the intrinsic motivation of improving ones' traffic safety, a raffle with a brand new car as the main prize

would be held for all participants in the study, to take place after the post-test collection (March 2004).

The introductory letter also announced that a follow-up telephone call would be made by RACC. This call was made in order to check the letter had been received, to remind recipients of its contents, address any questions and to confirm the participation or non-participation of the subject. If the answer was positive, a set of questions were asked to subjects during their telephone call, in order to collect some initial data on socio-demographic and driving experience variables.

The second mailing to the 621 participants left at this stage consisted of the pre-test, which was completed and returned by 350 subjects. According to the number and residence of the subjects in the test group, 7 one-day training sessions were organised, as follows: 3 in Valencia (July 4th, 5th and 6th), 3 in Barcelona (July 11th, 12th and 13th), and 1 in Madrid (July, 19th). A total of 137 subjects attended these training sessions that were organised by RACC according to the NovEV project guidelines. The control and test group participants that satisfied the previous steps received the third mailing with the post-test in December 2003, 263 subjects completed and returned the post-test.

After cleaning up the data and running some preliminary analyses, 25 of the subjects' data records were removed from the database because of missing or inconsistent response patterns. The final sample to consider in the evaluation was thus 238: 114 subjects from the control group and 124 from the test group.

RESULTS

The statistical analysis proceeded with some preliminary data analyses designed to check four basic aspects related to the validity of procedure: the characteristics of the sample of participants in relation to the reference population; the quality of the data collected; the metric properties of the measurement instruments; and the equivalence between the control and test groups at the outset of the research. Next, data analysis focused on the main goal of this study, namely to evaluate the effectiveness of the training course implemented by RACC according to the guidelines of CIECA's NovEV project. In this analysis, a double perspective was adopted according to the two sources of information collected from the participants in this research: driving behaviour indicators related to accident risk (criteria measures); and feedback on the training course they took (course evaluation measures).

Note that in the results presented in the following subsections an alpha level of .05 (denoting significant change) was used for all statistical analyses.

2.1. Preliminary data analysis results

2.1.1. Sampling- and self-selection- derived comparisons

Initial data analysis focused on checking how the sampling and self-selection processes had affected the characteristics of the sample of participants involved in the study. For this purpose, we compared the sample of subjects who agreed initially to participate in the experiment versus our reference population as described above in the 'sample' section. No significant differences should appear in the comparison of both groups in order to gain external validity of the experimental results, that is, people who participate in the experiment should not be significantly different from the original group of candidates (N=5728).

This comparative analysis (initial sample of participants v. experiment reference population) was possible because some variables were available from RACC database for most of the subjects on the initial list of candidates. Thus, the variables compared for these 2 groups were:

age; number of years as driving licence holder; number of accidents, car age; and car power. The gender variable was not considered in this comparison as it was blocked when randomly selecting the subjects who were posted and phoned with the invitation to participate. T-tests for independent samples were computed for each one of these numeric variables (see Appendix 1.1). Results showed that only for the variable 'Number of years having the driving licence', the difference between the mean in the population reference (M=1.85) and the mean for the initial sample of participants (M=1.74) was statistically significant ($t = 2.86$, $df = 5591$, $p = .005$). However, note that this difference (0.11) is equivalent to approximately one month and ten days only, and that the statistically significant difference could simply be motivated in this case by the large size of the samples compared.

2.1.2. Data quality control

The main focus of this section was the detection of errors and outliers in the dataset obtained after data collection and processing. This initial dataset contained 263 data records, each one corresponding to one of the participants who completed all the phases planned in the research design: selection questionnaire, pre-test, training session (only for the test group), and post-test. Four main strategies were applied to clean up the data collected: (1) anomalous values in frequency distributions, cross-tabulations, summary statistics, and uni- and bi- variate graphical representations; (2) subject's very low response rates; (3) extreme standardised means and standard deviations of the subject's response patterns (Dolinger & Dilalla, 1996), and (4) inconsistence of the subject's expected responses to specific pairs of items. While the first method is variable-oriented and provided information served either to correct some specific values or to re-code them as missing, the other 3 methods are case-oriented, providing support to locate strange response patterns, whether in the pre-test or in the post-test questionnaires. According to these criteria, data records from 25 subjects were removed from the database so the final sample size in the following analysis was 238 (see Appendix 2 for more details).

2.1.3. Psychometric analysis of the measures

This analysis was designed to validate the psychometric properties of the 5 scales used to collect the data that enabled us to test the impact of the training course (see Measures subsection). The five scales were analysed by obtaining the dimensional structure of the responses given to each scale, in order to test if the original factor structure was reproduced by our data. Dimensional analysis was carried out using the Principal Components Analysis (PCA) model to each one of the response data sets collected for the 5 scales in the pre-test questionnaire (where the number of subjects taking the scales was larger than in the post-test). The results of the factor extraction through PCA (see Appendix 3) showed that the scales A1, A2 and C2 had a mainly unidimensional factor structure, in line with the results of Hatakka (1998) and Keskinen et al. (1992) with Finnish samples of novice drivers. The Kaiser-Guttman criteria (eigenvalue > 1) was used to set the number of factors to be extracted, which resulted in 2 factors in the case of scales B and C1. However, this second factor in scale B explained a reduced amount of % of variance in relation to the first factor (15.4 v. 35.7), apart from having no sensible interpretation according to the item loadings in the second factor. With regard to scale C1, the mediocre KMO index for this scale (0,67) denotes that PCA is not very well suited to the response data collected for this scale; additionally, the second factor in the scale C1 appears as somewhat relevant if we look at its % of explained variance in relation to that of the first factor (17.8 v. 31.0). However, the low variance accounted for by this second factor (1.25) and the high saturation of all the items in the first factor suggest that it could be assumed that a main factor underlies the response data for this scale, so we cautiously considered it as unidimensional as well. As a consequence of these results, we worked henceforth with the subjects' scores corresponding to the factor scores associated with the first component extracted for each scale.

After analysis of dimensionality, we proceeded with the application of two classical psychometric procedures oriented to evaluate reliability of measures: Cronbach's alpha and the

test-retest indexes. The former was obtained for the data collected in the pre-test questionnaire because a larger sample was available than in the post-test application, and the results were: 0,73 for scale A1; 0.83 for scale A2; 0.68 for scale B; 0.62 for scale C1; and 0.81 for scale C2. Other results complementary to Cronbach's alpha are listed in Appendix 3.

With regard to the test-retest reliability index, this was estimated for each scale through the product-moment correlation between the scores obtained from the pre-test and post-test applications of the 5 scales. The calculation of the correlation coefficients was limited to the control group participants as, by not attending the training course, they received less external influences to change than the test group during the period between the pre-test and the post-test. The coefficients obtained were: 0.49 for scale A1; 0.41 for scale A2; 0.40 for scale B; 0.68 for scale C1; and 0.76 for scale C2, clearly higher for the driving habits self-report scales.

2.1.4. Equivalence of the test and control groups

A number of analyses were achieved to ensure that the experimental and control groups were equivalent at the outset of the research. With regard to the gender variable, as it was stated above, we blocked this variable in the random assignment of subjects to the research groups. However, the contingency table shows how the distribution of the 2 research groups by gender is far from the distribution of men and women in the initial sample of 621 participants: men: 51,7%; women: 48,3% (see Table 6). This was especially true for the control group (men: 40,4%; women: 59,6%), suggesting that the gender variable should be statistically controlled in successive data analysis process.

Table 6: Gender by Group contingency table

			Gender		
			Men	Women	Total
Research condition	Control Group	Count	46	68	114
		%	40,4%	59,6%	100,0%
	Test Group	Count	60	64	124
		%	48,4%	51,6%	100,0%
Total		Count	106	132	238
		%	44,5%	55,5%	100,0%

As was the case for the gender variable, the distribution of the test and control groups could no longer be as balanced as initially expected after the random assignment of the participants to the two research groups. Obviously, the loss of participants during the early phases of the project could have not followed a random path. So, checks were carried out to see to what extent the control and the test groups were balanced for some of the variables collected in the selection and pre-test questionnaires. More specifically, the checks covered the criteria (pre-test) and the control variables listed in the Measures subsection above.

In order to test if these variables were equally distributed in the control and test groups, we proceeded according to the type of variable:

- For the category variables, we obtained contingency tables for each one of these crossed by the variable Group. Pearson's chi-square test was used to contrast the independence of the distribution of the test and control groups with regard to the levels of each one of these variables. The results (see Appendix 1.2) showed that the control and the test groups were fairly balanced for all the variables considered: all the chi-square values had statistically non-significant results ($p > .05$). Only the variable Education (see Table 7) was on the limit of statistical significance ($\chi^2 = 5.98$, $df = 2$, $p = .050$), a possible explanation for it being

that working people (where supposedly more primary and secondary education levels are found) is less favourably disposed to attend a Saturday training programme.

Table 7: Education by Group contingency table

			Participants' educational level			Total
			Primary education	Secondary education	University education	
Research condition	Control Group	Count	32	48	32	112
		%	28,6%	42,9%	28,6%	100,0%
	Test Group	Count	25	43	53	121
		%	20,7%	35,5%	43,8%	100,0%
Total		Count	57	91	85	233
		%	24,5%	39,1%	36,5%	100,0%

- For the numeric variables, t-tests for independent samples were computed for each one of these variables with the research condition as the grouping variable. The null hypothesis of independence resulted in rejection of 6 of the 23 variables (see Appendix 1.2). Table 8 summarises this information for the 6 variables with significant differences.

Table 8: T-tests with statistically significant mean differences

	Research condition	N	Mean	t-test	Cohen d
Car power	Control Group	90	81,01	t (158) = 2.22 p = .028	0.360
	Test Group	70	73,99		
Mean scores for the 1st scale of training contents interests: Driving knowledge and skills	Control Group	113	3,4250	t (230) = -3.07 p = .002	0.406
	Test Group	119	3,6858		
Mean scores for the 2nd scale of training contents interests: Recognising and avoiding risks	Control Group	113	3,5630	t (229) = -2.07 p = .039	0.274
	Test Group	118	3,7624		
Mean scores for the 3rd scale of training contents interests: Recognising one's strengths and weaknesses	Control Group	113	3,9773	t (230) = -2.99 p = .003	0.418
	Test Group	119	4,2129		
Factor scores for the showing-off internal-risks scale (pre-test)	Control Group	114	-,2887	t (216) = -2.90 p = .004	0.395
	Test Group	119	,0923		
Factor scores for the skills for careful driving scale (pre-test)	Control Group	114	,1804	t (231) = 2.05 p = .042	0.269
	Test Group	119	-,0947		

If we look at the column with *Cohen's d* as estimation of the effect size, we can see that all the values are in the range of what is considered as small values of effect size (< 0.5), denoting that the statistically significant differences found in the analysis could be more related to a sample size factor than to the magnitude of the differences between the 2 research groups. Anyway, given that the effect sizes values for two of the training interests scales are over 0.4, it was advisable to consider these variables in successive analyses. For the sake of simplicity, taking into account that the 3 variables with the scores in the training interests scales work in the same way (having larger values for the test than for the control group), we assumed that a common factor of 'interest in improving driving' underlies all three, and we built a summary variable as a mean of them. This new variable would simplify later ANOVA modelling and, as it was expected, the test group (M = 3.89) reported larger 'interest in improving driving' than the control group (M = 3.65) according to the t-test computed for this new summary variable (t = -3.175, df = 230, p = .002, d = 0.419).

2.2. Main results

Results in this section focus on the specific statistical analysis achieved in order to assess the impact of the 2nd phase training for novice drivers carried out by RACC. This analysis has been achieved on a double basis: (1) the participants' scores in the five scales that were considered as the main dependent variables in this study; (2) the feedback obtained from participants after the course.

2.2.1. Scale-level results

An univariate ANOVA model was used to analyse the data of our mixed between-within design for each one of the five scales considered. In all cases, the Group condition (control v. test) was manipulated as a between-subjects factor, while the Time condition was manipulated as a within factor with also 2 levels: the repeated measures obtained for the scale in the pre-test and post-test moments. On the other hand, the variables Gender and the scores in the summary variable 'Interest in improving driving', which were non-equally distributed in the control and test groups, were also modelled in order to control their potential effect on the dependent variables.

The hypothesis about the impact of the training course would be supported if, first, the interaction Group by Time (or Group by Time by Gender) was statistically significant, and second, if the analysis of simple effects went in the expected direction: (1) differences between control and test groups in post-test scores; and (2) no differences between control and test groups in pre-test scores.

Prior analysis was carried out to check if the two main assumptions of the ANOVA method that must be upheld in the data were satisfied, namely the gaussian-shaped distribution of the criteria variables, and the homogeneity of the variances of the distributions defined by the levels of the independent variables. Different tests designed to check these two assumptions, like Shapiro-Wilk's and Levene's tests, showed that the normality and homogeneity of variance assumptions were, in general, not satisfied (see Appendix 4). Exploration of score distributions for the five scales appeared to be quite asymmetrically distributed, which makes sense if we take into account most of the item statements. An illustration of this fact is summarised in the scatter plot with the pre-test and post-test z-scores in the scale 'driving in an improper state' (see left figure below), which shows a non-homocedastic bivariate distribution that originates in two quite positively asymmetric distributions.

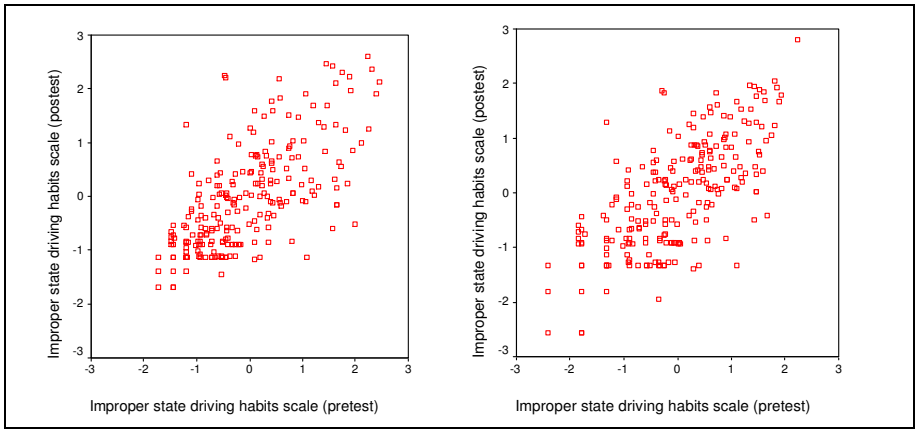


Figure 4: Pre-test and post-test z-scores in the scale 'Improper state driving habits' before (left) and after (right) data transformation.

In order to solve this problem, without losing the power of the ANOVA application in the analysis of our data, we opted for the strategy recommended in literature consisting of applying data transformations designed to obtain more symmetric, normal distributions which satisfy the

normality assumption and, simultaneously, improve the homogeneity of variances (Berry, 1987; Emerson, 1991).

According to the type and degree of asymmetry in the score distributions for the five scales considered in our study, we applied the transformation more suitable in each case: SQR or LN for the variables distributed with positive asymmetry; EXP(2) for those with negative asymmetry (see Appendix 4 for more details). The right plot in the figure above shows the same variables as in the left one, once a logarithm transformation was applied to the pre-test and post-test scores for this scale (and also a transformation to z-scores for easier interpretation and comparison). This new scatter plot shows how the transformation has modified the original asymmetry and spread of the cloud of the points. This visual observation was corroborated statistically through the acceptance of the hypothesis of normality and homogeneity of variances for these transformed two variables according to Shapiro-Wilk's and Levene's tests, respectively.

Thus, after transformation of the dependent variables in our study, an ANOVA mixed model was applied to the analysis of the data coming from each one of the 5 scales considered. Table 9 summarises the results (full lists in Appendix 5) showing only the F test value for the Group by Time interaction. No Gender by Group by Time significant interaction was found in any case so they are not included in Table 9.

Table 9: F-test for the interaction effect (Group by Time) for the 5 criteria measures.

SCALE	F	p
(Scale A1) Risks connected with personal careless driving habits	F(1,227) = 0.27	p = .60
(Scale A2) Risks connected with showing off and situational reactions	F(1,229) = 0.64	p = .42
(Scale B) Skills for careful driving	F(1,227) = 7.75	p = .01
(Scale C1) Driving habits related to being in an improper state	F(1,227) = 2.27	p = .13
(Scale C2) Driving habits related to high speed	F(1,227) = 0.13	p = .71

There were statistically significant effects only for the scale related to skills for careful driving so we proceeded with an analysis of simple effects. The results of this analysis (see Table 10) showed that the source of the significant interaction effect was the difference between the control and test groups in the pre-test moment.

Table 10: Pairwise comparisons (control v. test group) for the means derived from Scale B.

Pairwise Comparisons

Measure: MEASURE_1

time	(I) Research condition	(J) Research condition	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
1	Control Group	Test Group	,314*	,132	,018	,054	,573
	Test Group	Control Group	-,314*	,132	,018	-,573	-,054
2	Control Group	Test Group	-,079	,134	,557	-,343	,185
	Test Group	Control Group	,079	,134	,557	-,185	,343

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

a. Adjustment for multiple comparisons: Bonferroni.

In order to control this difference between the two groups at the outset of the research, we reanalysed the data taking the pre-test scores in this scale as covariate in the ANOVA model. The results of this ANCOVA analysis showed statistically significant differences between the control and test groups in post-test: F (1.228) = 5.07, p = .02 (see Table 11); therefore, empirically support was obtained for the effect of the course on the self-evaluation of skills for careful driving.

Table 11: ANOVA results for the 'Skills for careful driving' scale.

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	53,665(a)	4	13,416	16,967	,000
Intercept	,137	1	,137	,173	,678
group	4,006	1	4,006	5,066	,025
gender	,227	1	,227	,288	,592
FactorScor Pretest	48,022	1	48,022	60,730	,000
group * gender	2,199	1	2,199	2,781	,097
Error	180,290	228	,791		
Total	233,982	233			
Corrected Total	233,955	232			

Note that, although the gender by group interaction has a non-significant effect (p = .09), the graphic with the mean z-scores for the control and test groups split by gender (see figure below) shows how the differences in the means between the control and the test groups are more marked with the men than with the women.

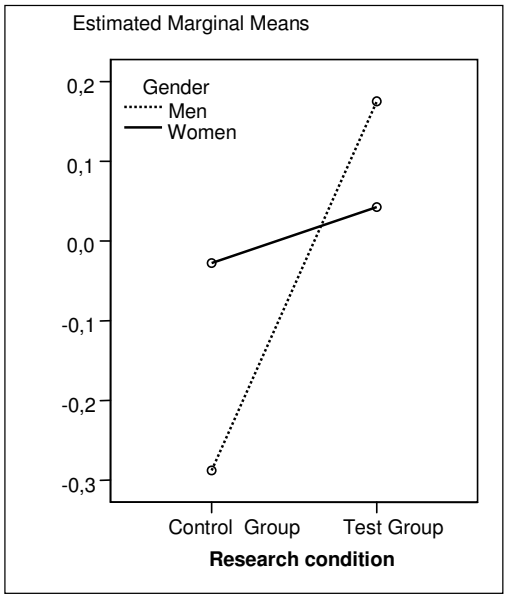


Figure 5: Mean z-scores for the control and test groups differentiated by gender.

2.2.2. Participants' post-training feedback

One hundred twenty participants from the test group answered the course feedback survey included in the post-test questionnaire. Table 12 shows the means and standard deviations for the scores in each one of the 10 items considered, which had to be answered for each of the three modules in the course (discussion-group, on-road and track training). Note that the scale

ranged from 1 (Not at all) to 5 (Very much), and how the item 4 ('It was rather boring') was reversed in order to measure this in the same way as the rest of the items. Thus, the descriptive statistics stand for 'interesting' rather than 'boring'.

Table 12: Descriptive statistics for the items in the course evaluation questionnaire

	On road training		Track training		Discussion group		
ITEMS	**Mean**	**St. Dv.**	**Mean**	**St. Dv.**	**Mean**	**St. Dv.**	**MEAN**
10. I recommend this training for young drivers.	4,51	,83	4,85	,46	4,97	,22	4,77
3. The trainer(s) did a good job.	4,47	,78	4,66	,56	4,81	,42	4,64
4. It was rather boring (REVERSED).	4,32	,93	4,76	,69	4,58	,87	4,55
2. It was well organized.	4,32	,68	4,53	,56	4,67	,51	4,50
9. It was interactive: we had opportunities to question and discuss	4,28	,83	4,26	,90	4,47	,75	4,34
5. It helped to me to become more aware of certain risks when...	3,48	1,11	4,23	,84	4,48	,82	4,06
1. I learnt some interesting things related to driving.	3,15	1,14	4,46	,77	4,46	,73	4,02
7. It made me reflect on some bad driving habits I had.	3,47	1,12	3,69	1,15	4,25	1,04	3,80
6. I learnt some driving techniques to better control the car.	2,48	1,12	4,02	1,02	2,71	1,27	3,07
8. It helped me become a more skilful driver.	2,50	1,15	3,41	1,23	2,68	1,30	2,86
MEAN:	3,70		4,29		4,21		

Item scores for the 10 items in the 3 courses were modelled with an ANOVA mixed model with 2 within-variables ('Items' with 10 levels and 'Course' with 3 levels) and the Gender variable as between-variable. Given that the latter had no major effect, nor participated with the other variables in significant interaction effects, it was removed from the definite model applied to analyse the data.

Table 13: Simple effects analysis for the Item by Course interaction.

ITEMS	On-road training v. Track training	On-road training v. Group discussion	Track training v. Group discussion
1. I learnt some interesting things related to driving.	P < 0.01	P < 0.01	n. s.
2. It was well organized.	P < 0.01	P < 0.01	P < 0.01
3. The trainer(s) did a good job.	P < 0.05	P < 0.01	P < 0.05
4. It was rather ~~boring~~ interesting.	P < 0.01	P < 0.05	n. s.
5. It helped to me to become more aware of certain risks when driving.	P < 0.01	P < 0.01	P < 0.05
6. I learnt some driving techniques to better control the car.	P < 0.01	P < 0.05	P < 0.01
7. It made me reflect on some bad driving habits I had.	n. s.	P < 0.01	P < 0.01
8. It helped me become a more skilful driver.	P < 0.01	n. s.	P < 0.01
9. It was interactive; we had opportunities to question and discuss	n. s.	P < 0.05	P < 0.05
10. I recommend this training for young drivers	P < 0.01	P < 0.01	P < 0.01
TOTAL (Mean Scores):	P < 0.01	P < 0.01	n. s.

The ANOVA results showed statistically significant main effects for the Item and the Course factors as well as for the Item by Course interaction, so an analysis of simple effects was computed. The results of this analysis (see Table 13) showed that the source of the significant interaction effect is associated to different significant mean differences between the three parts of the course. Many of them are motivated by significant lower scores in the items for the on-road part than for the track and group discussion parts.

A graphical visualisation for these results is shown in Figure 6, which allows for an easier interpretation of these differences. From these results, a rather positive evaluation was reported by the participants for the 3 parts of the training course in the items where it would be expected (mean scores over 3.80). This is specially true for the course organization and the trainers' work items, as well as with regard to 'becoming more aware of certain risks when driving'. On the other hand, lower average scores were obtained for the items related to learning driving techniques to better control the car (mean score: 3.07) and to becoming a more skilful driver (mean score: 2.86). Anyway, some specific results provide indications about some considerations to take into account in the implementations of future training courses: (1) Items

#5 and #6 show significant lower scores for the on-road training, so it might be advisable to revise the contents and follow-up with more discussion of the on-road exercises in order to improve this module; (2) items #6 and #8 score significantly higher for the track training, which may indicate an overestimation of driving skills - something that should be avoided in the potential implementation of the course in future. Additionally, a more detailed questionnaire is desirable for collecting the participants' feedback in order to know the real reasons behind some of the general statements of the items in our scale.

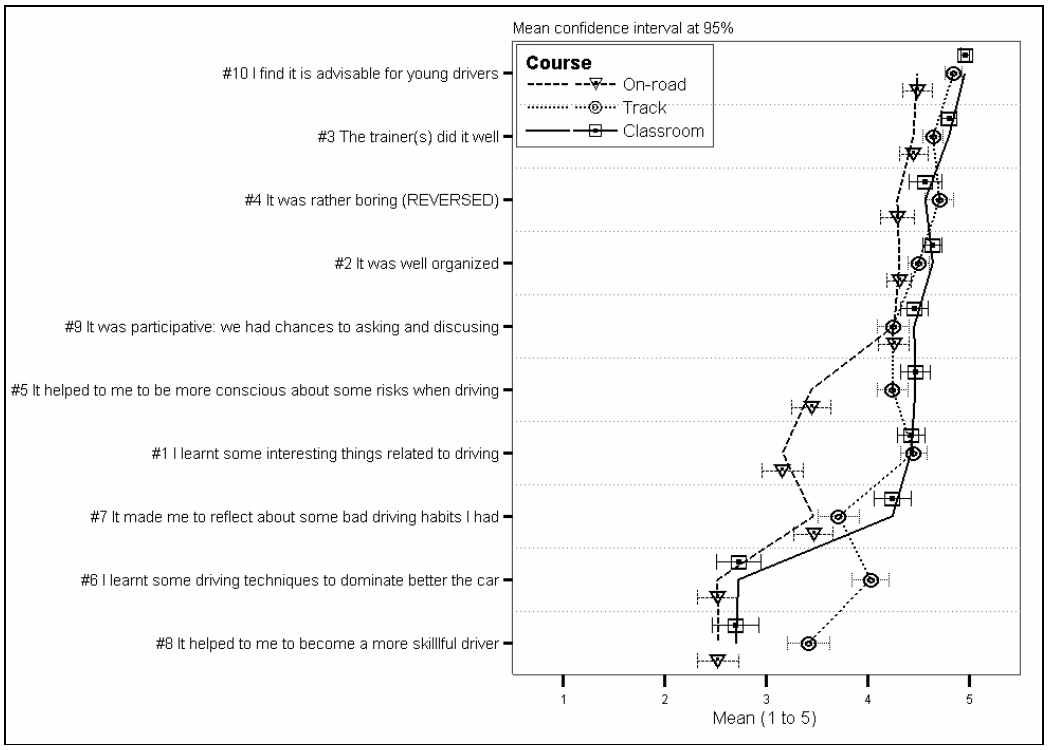


Figure 6: Item mean scores for the three training parts in the course.

3. Conclusions

The main conclusions of the RACC pilot trial of the NovEV project are positive if the findings and results from the training are compared to the current system of learner driver education in Spain. The implementation of this project has proved the following:

1. **Training results**: Data analysis results for the criteria measures considered in this study show statistically significant differences between the control and test groups for the "Skills for Careful Driving" scale, meaning that the mean score in this scale was higher for the test group than for the control group after the training. This result goes in the expected direction given that, as reported in earlier studies, self-evaluation of skills for careful driving is inversely related to accidents.
2. **Research design**: Positive differences between the test and control groups were also found for the other four driving behaviour scales, but these differences did not appear to be statistically significant, so they may have occurred by chance. Perhaps, significant differences could have arisen if the research design had been improved, such as: (1) a larger sample size; (2) a more balanced sample selection -in our study the control group appeared in the pre-test as systematically 'safer' than the test group; and (3) scales with items which were better tailored to the changes expected as a result of each course module. As some items were perhaps not very directly related to the course subjects, this may have hampered our attempts to detect specific changes.
3. **Participant's feedback**: The second line of data analysis, the one based on the participants' feedback, offered interesting input for evaluating the course itself and the effects of the course as reported by the participants. With regard to the former, mean scores over 4 (in a scale from 1 to 5) for the items related to the course organisation, contents, and tuition revealed a rather positive evaluation for the course.
4. **Self-reported improvements**: With regard to the improvements reported by the participants as a result of the course, it appeared that the mean scores were significantly higher for the items related to self-awareness about risks and bad driving habits than for the items on driving techniques and skills. These results conform with the aims of the NovEV project for such post-licence training, and this bodes well for potential implementation in the future.
5. **Training implementation**: the priority was to implement a short and feasible training, rather than a long and therefore more comprehensive - but also less feasible - training. However, it seems that the training sessions (track, road and workshop) should be extended from 90 minutes to 120 minutes in order to generate better feedback discussions between participants and trainers in each session.
6. **Further improvements**: a separate comparison of the evaluation of the 3 course modules showed that more efforts should be devoted to improving the on-road part of the training, which was often perceived as a typical driving lesson rather than a feedback drive. Therefore, training-the-trainer must be reinforced in this part and the observation forms have to be simpler. The track training is well designed, but needs a better feedback session allowing trainees to raise questions and issues by themselves. Finally, the workshop session should encourage trainees to participate more actively by providing their own experiences. A short break should also be added in the middle.

REFERENCES

Berry, D. A. (1987). Logarithmic transformations in ANOVA. *Biometrics, 42,* 439-456.

Dolinger, S. J. & Dilalla, D. L. (1996). Cleaning up data and running preliminary analyses. In Frederick T. L. Leong & James T. Austin (Eds.): *The Psychology Research Handbook. A Guide for Graduate Students and Research Assistants.* Thousand Oaks: Sage

Emerson, J. D. (1991). Introduction to transformations. In *Fundamentals of Exploratory Analysis of Variance.* D. C. Hoalglin, F. Mosteller, & J. W. Tukey (Eds.). NY: Wiley & Sons.

Hatakka, M. (1998). Novice Driver's Risk- and Self-Evaluations (Doctoral Thesis). Annales Universitatis Turkuensis, ser.B- TOM.228. Turku.

- SPAIN RACE

NovEV
Results of the RACE second phase

Tomas Santa-Cecilia, Javier Sanchez

1.1. Study Summary

The general target of this study was to evaluate scientifically the influence of second phase training courses on novice drivers (once the driving licence has been obtained and some driving experience accrued) and to measure any changes related to skill, knowledge, behaviour and attitude as a result of the course.

The sample participating in the study was composed of 154 subjects. This sample was selected according to the following selection criteria: category B licence holders for between one/two years and a minimum of 5.000 km of driving experience. Once selected, the participants were randomly assigned to two groups:

- Experimental group (77 participants)
- Control group (77 participants)

The experimental group took part in the training. The course contents were focused on a few very clear messages, especially oriented to road safety. The Training Programme was composed of three different modules: class (theoretical contents, discussion), track (practical contents), real traffic (assessed driving). The control group did not participate in the training. The aim of the control group was to establish the base line in order to determine which part of the change achieved in the experimental group was due to the training and which part was due to the driver's natural development.

Evaluations were conducted at 3 stages with a view to establishing the differences between the two groups over a period of six months. The first evaluation (pretest) took place before the training programme in order to establish a base line and to be able to compare later evaluations. Two further evaluations followed after the training programme: after a week (to evaluate any results over the short-term) and after six months (to evaluate results over the medium to long term).

The following data collection methods were used for this purpose: a road safety questionnaire and an evaluation in real traffic (driving assessment on public roads) in order to compile as much data as possible related to current knowledge, skills, behaviour and attitudes of the participants. The data analysis methods used in the study were a descriptive analysis and ANOVA.

In the on-road evaluation, a significant improvement in general driving skills within the experimental group was found as a result of the training. In the attitude variable, no differences were found. According to the questionnaire feedback, there was a significant improvement in knowledge within the experimental group as a result of the training.

CONCLUSIONS:

Analysing the results obtained, we can conclude that there was a significantly higher change in knowledge, skills and behaviour (in the attitude variable the results are not so favourable as expected, so no conclusions can be drawn from these) in the experimental (training) group than in the control group. Therefore, the training was seen to have a considerable effect on participants in the short and medium term (6 months). We can thus conclude that there was a positive effect of the course on novice drivers.

In the skills and behaviour variables we found that:
The training improved the participants' driving skills and behaviour in comparison to the control group
In the attitude variable we found that:
No differences were found between the results obtained in the experimental group and the control group
In the knowledge variable we found that:
The training improved the road safety knowledge of the participants in comparison to the control group.

1.2. Participants

The initial aim was to obtain a sample of 396 participants. More than 5.000 letters were sent by pre-selected driving schools to pupils who complied with the selection criteria. However, only 216 filled in the recruitment form.
So, the initial sample of our study was composed of 216 subjects. Among these, only 154 completed the whole process, culminating in a drop-out rate of 29%.

Selection criteria of the sample:

Only drivers with a minimum of 5.000 km of experience and holding a category B driving licence for between one/two years were eligible to participate in the study.

To select the participants, a recruitment form was distributed to be filled in by the novice drivers, in which the following information was collected in order to divide the participants and to carry out the statistical work:

-Personal data
- Age
- Sex
- Current residence: town or village
- New drivers or other previous licences
- Number of attempts at the theory or practical tests
- Length of education
- Normal use of car
- Accidents, incidents and traffic offences

Once the participants were selected, they were randomly assigned to one of two groups.
The two groups were:

- The Experimental group, participating in the training (77 participants)
- The Control group, not participating in the training, whose purpose was to establish the base line in order to determine which part of the change achieved in the experimental group was due to the training and which part was due to the driver's natural development (77 participants).

In order to avoid any selection bias, the participants were totally unaware of the selection criteria. On a similar note, there was a risk that the subjects taking part in the project had an above-average positive attitude to road safety and good traffic behaviour, and were not therefore representative of the general population. To avoid this, material incentives were offered to the participants in order to motivate them to take part (a lottery to win a Renault Clio, gifts such as jackets, books, etc.)

1.3. Training programme

It is important to bear in mind the difficulty of designing and executing an effective one-day training programme, which results in positive changes amongst participants. With this in mind, the course focused on a small number of very clear messages oriented towards road safety and self-awareness. Often, too much information is offered and little is retained by the novice driver. For this reason, we decided to focus on quality, rather than quantity.

The information was presented in an interesting way in order to maintain the participants' attention, to encourage self-expression (opinions, comments and changes of ideas and experiences), by taking an active role during the whole training programme via group dynamics, self-analysis and guided discussion.

In this sense, the trainers were essential for establishing a rapport with the participants and for transmitting the concepts and objectives of the training. To make the training more homogeneous, the courses were led by the same trainer in all the courses in the four cities.

The task of this Road Safety Trainer was to impart course aspects related to knowledge acquisition, by supporting the psychologist's work, whose aim was to influence attitudes and behaviour with regard to Road Safety. To sum up at the end of the theoretical training, one person belonging to AESLEME (a Spanish NGO raising awareness on spinal and brain injuries) raised the issue of the consequences of an accident.

Design of the training course

The training was composed of 3 modules:

Place	**Nº. Hours**	**Trainers**
Classroom	4	1 Road Safety Trainer 1 Psychologist AESLEME
LUNCH		
Track	2	1 Teacher 1 Psychologist
Real Traffic	1	1 Road Safety Teacher
Discussion and close	1	1 Road Safety Teacher 1 Psychologist

A) Design of each training module in class:

a) Each training module started with a group discussion in order to assess the experiences, opinions, strong and weak points and self-evaluation of each participant.
b) The trainer explained the basic concepts of the training module by introducing information with attractive documentation and terminology which was easy to understand for the participant.
c) Again, a group discussion took place on the participants' opinions and discussion on the subjects addressed.
d) At the end of the training in class, a summary was made of the concepts, ideas and main messages transmitted during the different training modules which the course consists of.

The participants' learning potential is greater when he/she is refreshed enough, so a short break of approximately 5-10 minutes was held every one and half hours. Here the participant could get up, breath some fresh air, drink something and chat with other participants and the trainer.

- Contents of the training in class:
  - Data about accidentology
  - Perception of risk
  - Lapses in concentration
  - Speed and its relation to accidents
  - Objects inside the vehicle
  - Physical and Psychological state
  - Alcohol, Drugs and their consequences
  - Security Features
  - Other road users
  - Effect of Age, of young people between 18 and 24 years old
  - What to do in case of accident

B) Lunch

During the lunch hour, the training continued in a more relaxed atmosphere, where the participants were more open to share and change their experiences and interests. The idea was to reach the ideal social atmosphere to empower self-analysis, communication, relaxation and the right diet.

C) Design of each track training module

Standing on the track, the trainer explained the importance of the steering-wheel position and other safety features in a stationary vehicle.

Before the beginning of each exercise, the trainer made a demonstration of the exercise, explaining its difficulty and its objective. The exercise was repeated by the driver up to a maximum of three times to avoid over-confidence, and so the pupil could become aware of his/her limits and the limitation of the vehicle.

D) Design of each training module in real traffic:

In groups of three and with a trainer, the drivers had to drive along a route with the following features:

- Incorporation to motorway / main road
- Exit of motorway / main road

- Lane selection
- Overtaking
- Driving on regional roads
- Driving around roundabouts
- Junctions / Intersections
- Driving around bends

The aspects addressed in the real traffic module were:

- Safety features
- Frequent mistakes
- Defensive driving
- Ecological driving
- Aggressive driving
- Wrong speed
- Safety distances
- Other road users

E) Design of the closing session of the training course

The aim of the final hour of the training was to clarify any doubts or questions which may have arisen over the course, to review the main concepts and messages, by encouraging the active participation of the participants to reach the desired conclusions. This phase is essential for consolidating memory retention and for learning over a longer period. The idea of the "End Message" is crucial.

1.4. Trainers

The cooperation of certain driving schools was required. Each one assigned an instructor who took part in the training in real traffic, made the evaluation and filled in the evaluation questionnaire during this module. For the theoretical/group discussion part and the track training in track, RACE assigned an expert in these subjects, teacher D. Jorge Castellanos (technical expert).
There was no training of these trainers, but there were some workshops to homogenise evaluation criteria.

1.5. Quality evaluation

These are the results of the satisfaction questionnaire from the training day:

Quality evaluation Questionnaire	**Mean**	**Std**
Teachers level	7,5	0,8
Course contents	7,9	1,1
Didactic Material	7,7	0,9
Development of the theoretical course	7,4	0,7
Development of the practical course	7,8	1,5
Training in real traffic	8,2	1,3
GLOBAL EVALUATION	7,7	0,8

Scale 1-10 (1= poor, 10= excellent)

In addition, almost 90% of the course participants were satisfied with the training and 95% of the participants would recommend the training to others.

1.6. Evaluation Design

People responsible for the evaluation
The people assigned to making the practical evaluation were driving instructors employed in the driving schools participating in the project. They had to fill in an assessment form once the driver had finished the on-road driving assessment.

In addition to the experimental group, the control group also received the same evaluation system.

1) What was evaluated / measured
Each measurement of the participant corresponded to the effects of the training on different levels: knowledge, skills, behaviour and attitudes.

Knowledge
- Data about accidentology
- Speed and its relation to accidents
- Objects inside the vehicle
- Alcohol and its consequences
- Safety features
- Other road users
- What to do in case of accident

Skills:
- Steering-wheel position
- Taking bends

Behaviour:
- Most frequent mistakes
- Risk perception
- Wrong speed
- Safety distances
- Inattention
- Physical and psychological state
- Alcohol, drugs
- Other road users
- Laws relating to safety features

Attitudes:
- Defensive driving
- Ecological driving
- Age-related factors= young people between 18 and 24 years old
- Aggressive driving

2) When and how was the evaluation carried out?
The evaluation was carried out as follows:
1. Prior to training: to know the current state of the participants in the training and control groups.
2. At the end of the training day: to know the satisfaction level of the pupil with the training course.
3. One week after finishing the training: to know the state of the participants in the training group and to compare it with the results received in the evaluation in point 1.

4. 6 months after having finished the training: to know the state of the participants in the training and control groups and to compare the results already received.

1. **Prior to training:**

a) Questionnaire

The aim was to collect information relating to knowledge, attitudes, and driver behaviour

b) Evaluation in real traffic

The aim was to collect information related to current knowledge, attitudes and behaviour of the participant. This is the most effective measuring method, namely by direct observation, establishing a route of approximately 40 minutes duration covering situations we wanted to evaluate in town, on national roads and on the motorway.

2. **Immediately after completion of the training**

a) Questionnaire on training course quality

Its aim was to collect information on the satisfaction level of the pupil in all aspects of the course, the knowledge transmitted, the teachers, material, time distribution and all other aspects related to quality.

3. **One week after the training**

Its aim was to measure how much and how the pupil has changed due to his participation in the course, in the very short-term. This evaluation was only carried out with regard to the experimental (training) group, because the participants in the control group were evaluated before and they would not present significant changes due to only one week of additional experience.

4. **Six months after completion of the training**

The aim of this evaluation was to measure over the medium/long term any changes occurring amongst participants in the training and control groups.

1.7. Data collection methods

1. Evaluation Questionnaire

The aim here was to collect information related to current knowledge, behaviour and attitudes of the participant. The questions were aimed at assessing the level of knowledge in areas related to road safety, opinions, thoughts, intentions (attitudes) and driving style, including habits (behaviours). This questionnaire was filled in by the pupils in the initial phase of the study (see annex 12).

2. Evaluation in real traffic

Here, information was collected with regard to current attitudes and behaviour of the participant. This is the most effective measuring method, namely by direct observation, establishing a route of approximately 40 minutes duration covering situations we wanted to evaluate in town, on national roads and on the motorway. For this, an assessment form was designed, to be filled in by the on-road evaluator (see annex 12).

3. Questionnaire of quality training

The satisfaction levels of the participants were collected in relation to the course, the knowledge transmitted, the teachers, material, time distribution and all other aspects related to quality (see annex 12).

1.8. Methods used in the statistical analysis

For the general data analysis, descriptive statistics were used based on means and standard deviation. The evaluation questionnaires were tested for internal consistency using Cronbach's alpha based on the correlation interelements promedium.

All the tests were correlated using the quantitative focus bivariated of Pearson.
For the contrast of hypothesis, the following models were used:

- ANOVA of repeated measures with two factors:
  - Within design: evolution of the subjects, by means of three measures in the experimental group (before the course, after the course and six months later) and two measures in the control group (before the course and six months later).
  - Between design: defined group with two levels, experimental and control groups.

- ANCOVA of one factor, defined the analysis in the following way:
  - dependant variable: the total mark of the subjects once finished the course.
  - factor as the group: experimental and control
  - covariable, the initial mark

The aim was to measure the effects of the training course controlling the possible initial differences between the experimental and control group.

Tukey and Bonferrony were used to check the possible post-training differences between the means.

2. Evaluation of results

2.1. On-Road Evaluation (driving skills and behaviour)

The aim of the on road evaluation was to collect information related to current skills and behaviour of the participants for to determine whether this post-licence driver training courses for novice drivers are effective and have a positive effect. These questionnaires were studied and underwent reliability and validity tests, getting acceptable and good results (α de cronbach=0.95, see annex 12, separate document; analysis for the rest of variables can be found in separate document, available from CIECA on request).

The score in the on road evaluation was obtained by adding each of the variables included in the test: 1. Accommodation and handling (adjustment, seat belt and use of controls), 2. Circulation, 3. Safety Distance 4. Speed. The scores oscillate between 0 and 10, 0 being the minimum score and 10 being the maximum score. A high score in the test would indicate a high level of knowledge, skills, aptitudes and attitudes that the participant shows in the on-road evaluation.

Table 2.1.1 variable score on road evaluation * group

GROUP		INISCORE	WKSCORE	SIXSCORE
Experimental Group	Mean	4,53	5,84	5,55
	N	74	77	76
	Std. Deviation	1,30	1,22	1,45
Control Group	Mean	4,53		4,57
	N	75		73
	Std. Deviation	1,49		1,41
Total	Mean	4,53	5,84	5,07
	N	149	77	149
	Std. Deviation	1,40	1,22	1,51

INISCORE: 1ª evaluation score variable total score
WKSCORE: After the course score variable total score
SIXSCORE: After 6 months score variable total score

At a descriptive level, we can observe differences within the experimental group (within groups) and differences between the experimental group and the control group (between groups).

Within the experimental group, we can observe that the lowest scores coincide with the first evaluation (pre-test) and the highest scores coincide with the evaluations realized after applying the training course (after one week and to the 6 months).

The analysis of the scores of the control group allows us to study the effect of the training course on the experimental group. Within the control group, we observe that there are hardly any differences between the results obtained in the pre-test and in the post-test.

As a result of this data, we can conclude that there is a short term positive effect of the training course and, indeed, a long term lasting effect, since the effects remain over time. The results also allow us to conclude that the effect is due to the training course and not as a result of learning outside of the course when just driving, since no improvement was observed in the control group.

ANALYSIS BY GENDER

Table 2.1.2 Analysis by gender in on-road evaluation

GENDER		INISCORE	WKSCORE	SIXSCORE
men	Mean	4,56	5,85	5,11
	Std Deviation	1,38	1,17	1,52
women	Mean	4,47	5,82	4,98
	Std Deviation	1,43	1,32	1,50
total	Mean	4,53	5,84	5,06
	Std Deviation	1,39	1,22	1,50

Table 2.1.3 Analysis by gender in on-road evaluation * group

GENDER	GROUP	INISCORE	WKSCORE	SIXSCORE
men	Experimental group			
	Mean	4,57	5,85	5,53
	Std Deviation	1,16	1,17	1,54
	Control group			
	Mean	4,55		4,66
	Std Deviation	1,58		1,37
	Total Mean	4,56	5,85	5,11
	Std Deviation	1,38	1,17	1,52
women	Experimental group			
	Mean	4,45	5,82	5,59
	Std Deviation	1,56	1,32	1,29
	Control group			
	Mean	4,49		4,38
	Std Deviation	1,34		1,47
	Total Mean	4,47	5,82	4,98
	Std Deviation	1,43	1,32	1,50
total	Experimental group			
	Mean	4,53	5,84	5,55
	Std Deviation	1,30	1,22	1,45
	Control group			
	Mean	4,53		4,56
	Std Deviation	1,49		1,40
	Total Mean	4,53	5,84	5,06
	Std Deviation	1,39	1,22	1,50

As we can observe in the results, not much difference exists in the means between men and women. The scores are quite high amongst males. The highest scores coincide with the evaluations after the training (after one week).

ANALYSIS OF RESULTS FROM ON ROAD ASSESSMENT

Table 2.1.4 ANOVA results of the change variable (driving skills and behaviour)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
CHANGE	17,97	1	17,97	18,47	,00
CHANGE * GROUP	23,40	1	23,40	24,04	,00
Error(CHANGE)	138,20	142	,97		

The change variable= differences between iniscore variable(1) and sixscore variable(2)

Table 2.1.5 ANOVA Comparisons between experimental group and control group

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Experimental	Group	,46*	,20	,02	6,32E-02	,87
Control Group		-,46*	,20	,02	-,87	-6,32E-02

Differences were found between the results obtained in the control group and the results obtained in the experimental group ($p<0.05$).

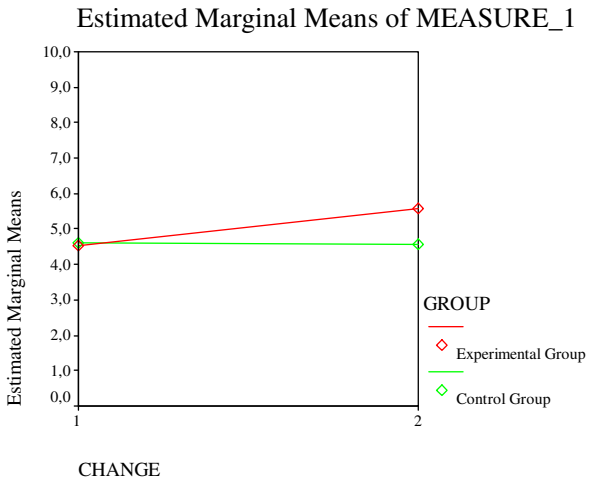
Table 2.1.6 ANOVA Comparisons between iniscore variable (1) and sixscore variable (2)

(I) CHANGE	(J) CHANGE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-,49	,11	,00	-,73	,27
2	1	,49	,11	,00	,27	,73

(1) score obtained in the 1ª evaluation, with regard to the skills and behaviour of the participants
(2) score obtained after 6 months, with regard to the skills and behaviour of the participants

Differences were found between the results obtained in the 1ª evaluation and the results obtained after 6 months (($p<0.05$).

Profile Plots



ANALYSIS OF THE EXPERIMENTAL GROUP (3 MEASURES)

Table 2.1.7 ANOVA results only experimental group (driving skills and behaviour)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
CHANGE	69,49	2	34,74	40,81	,00
Error(CHANGE)	122,57	144	,85		

The change variable= differences between iniscore variable(1), wkscore variable(2) and sixscore variable (3)

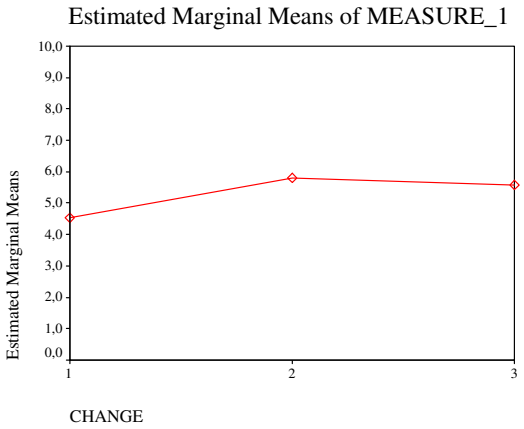
Table 2.1.8 ANOVA Comparisons between iniscore variable(1), wkscore variable(2) and sixscore variable(3)

(I) CHANGE	(J) CHANGE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-1,28*	,14	,00	-1,58	-,99
	3	-1,06*	,15	,00	-1,36	-,77
2	3	,21	,16	,17	-,10	,54
	1	1,28	,14	,00	,99	1,58
3	2	-,21	,16	,17	-,54	,10
	1	1,06	,15	,00	,77	1,36

(1) score obtained in the 1ª evaluation, relate with the skills and behaviour of the participants
(2) score obtained after the course, relate with the skills and behaviour of the participants
(3) score obtained after 6 months, relate with the skills and behaviour of the participants

Differences were found between the results obtained in the 1ª evaluation and the results obtained after the course, and the results obtained after 6 months ($p<0.05$), analysing only the experimental group.
We have not found differences between the results obtained a week after carrying out the course and the results obtained after 6 months ($p>0.05$)

Profile Plots



ANALYSIS OF THE CONTROL GROUP (2 MEASURES)

Table 2.1.9 ANOVA results only control group (driving skills and behaviour)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
CHANGE	,17	1	,17	,15	,69
Error(CHANGE)	79,02	70	1,12		

The change variable= differences between iniscore variable(1) and sixscore variable(2)

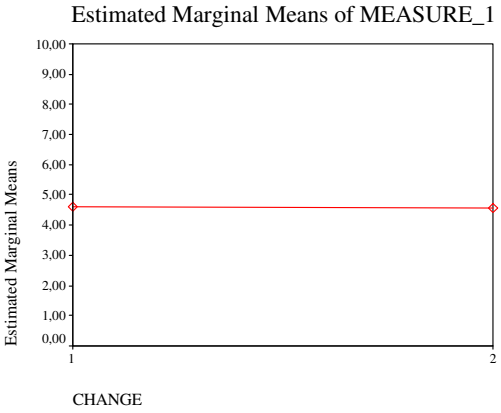
Table 2.1.10 ANOVA Comparisons between iniscore variable(1) and sixscore variable(2)

(I) CHANGE	(J) CHANGE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	7,04E-02	,17	,69	-,28	,42
2	1	-7,04E-02	,17	,69	-,42	,28

(1) score obtained in the 1ª evaluation, relate with the skills and behaviour of the participants
(2) score obtained after 6 months, relate with the skills and behaviour of the participants

No differences were found between the results obtained in the 1ª evaluation and the results obtained after 6 months ($p>0.05$), analysing only the control group.

Profile Plots



ANCOVA
VARIABLE DEPENDENT: sixscore
FIXED FACTOR: group
CO-VARIABLE: iniscore

Table 2.1.11 ANCOVA results of the SIXSCORE VARIABLE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	116,08[b]	2	58,04	37,42	,00
Intercept	80,94	1	80,94	52,18	,00
INISCORE	77,36	1	77,36	49,87	,00
GROUP	42,90	1	42,90	27,66	,00
Error	218,68	141	1,55		
Total	4049,46	144			
Corrected Total	334,77	143			

Conclusions: On-road assessment

Within the comparisons carried out by the on-road evaluation we find that:
Significant differences exist:
- Between the results obtained in the control group and the results obtained in the experimental group.
- Within the experimental group, significant differences exist between the results of the 1ª evaluation and the results obtained a week after the course and after 6 months.

Significant differences do not exist:
- Within the experimental group, between the results obtained a week after carrying out the course and the results obtained after 6 months.
- Between the results of the control group in the initial evaluation and the control group after 6 months.

2.2 Evaluation of results: Road Safety Questionnaire

2.2.1 Attitude

To collect the data several questionnaires were created with a view to collecting as much information as possible on the attitudes of the participants. These questionnaires were studied and underwent reliability and validity tests, obtaining acceptable results (α de cronbach=0.59, see annex, separate document, available on request).

The scores in the attitude variable have been obtained by adding the scores of each of the subjects in the questions that are focused on attitudes in the questionnaire. The scores range between 0 and 10, 0 being the lowest score and 10 being the highest. A high score in the attitude variable would indicate a predisposition towards safe and responsible driving. Scores were based on a series of questions related to factors such as:

- Causality of accidents
- Risk perception
- Perception of safe distances
- Strategic planning for driving
- Speed
- Use of items that can distract / affect one's concentration.
- Sleep and alcohol

Next, the following descriptive data is presented (means and standard deviation) of the variable attitude, based on an analysis of groups.

Table 2.2.1.1 Variable ATTITUDE * GROUP

GROUP		INIATTIT	WKATTIT	SIXATTIT
Experimental Group	Mean	7,49	8,11	7,93
	N	66	63	63
	Std. Deviation	,84	,87	,74
Control Group	Mean	7,74		7,52
	N	71		71
	Std. Deviation	,69		,80
Total	Mean	7,62	8,11	7,72
	N	137	63	134
	Std. Deviation	,77	,87	,80

INIATTIT: 1 evaluation score variable attitude
WKATTIT: After the course score variable attitude
SIXATTIT: After 6 months score variable attitude

At a descriptive level, we can observe differences within the experimental group (within groups) and differences between the experimental group and the control group (between groups). Within the experimental group, we can observe that the lowest scores coincide with the first evaluation (pre-test) and the highest scores coincide with the evaluations realized after the training course took place (after one week and to the 6 months).
Within the control group, we observe that there are hardly any differences between the results obtained in the pre-test and in the post-test.

ANALYSIS BY GENDER

Table 2.2.1.2 Analysis by gender ATTITUDE variable

GENDER		INIATTIT	WKATTIT	SIXATTIT
men	Mean	7,68	8,18	7,77
	Std Deviation	,77	,90	,81
women	Mean	7,51	7,95	7,61
	Std Deviation	,76	,79	,77
total	Mean	7,62	8,10	7,71
	Std Deviation	,77	,87	,80

Table 2.2.1.3 Analysis by gender * group ATTITUDE variable

GROUP	GENDER		INIATTIT	WKATTIT	SIXATTIT
Experimental Group	Men	Mean	7,62	8,18	8,03
		Std Deviation	,83	,90	,68
	women	Mean	7,24	7,95	7,75
		Std Deviation	,80	,79	,82
	Total	Mean	7,49	8,10	7,93
		Std Deviation	,83	,87	,74
Control Group	Men	Mean	7,73		7,54
		Std Deviation	,72		,86
	women	Mean	7,75		7,49
		Std Deviation	,64		,71
	Total	Mean	7,74		7,52
		Std Deviation	,69		,80
Total	Men	Mean	7,68	8,18	7,77
		Std Deviation	,77	,90	,81
	women	Mean	7,51	7,95	7,61
		Std Deviation	,76	,79	,77
	Total	Mean	7,62	8,10	7,71
		Std Deviation	,77	,87	,80

As we can observe in the results, not much difference exists between the mean scores of men and women. The scores are quite high amongst males. The highest scores coincide with the evaluations after the training (after one week).

Analysis of variance of the variable ATTITUDE

ANOVA
ATTITUDE

Table 2.2.1.4 ANOVA results of the change variable (attitude)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
CHANGE	,36	1	,36	1,06	,30
CHANGE * GROUP	5,88	1	5,88	17,10	,00
Error(CHANGE)	40,94	119	,34		

The change variable= differences between iniattit variable(1) and sixattit variable(2)

Table 2.2.1.5 ANOVA Comparisons between experimental group and control group (attitude)

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Experimental	Group	4,53E-02	,11	,70	-,19	,28
Control Group		-4,53E-02	,11	,70	-,28	,19

No differences were found between the results obtained in the control group and results obtained in the experimental group (p>0.05) in the attitude variable.

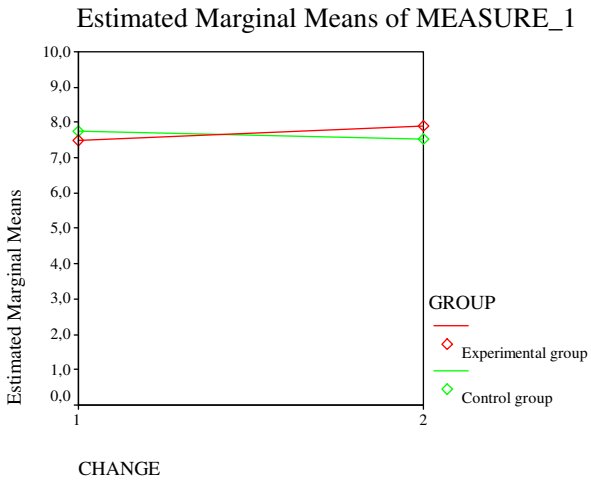
Table 2.2.1.6 ANOVA Comparisons between iniattit variable (1) and sixattit variable (2)

(I) CHANGE	(J) CHANGE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-7,84E-02	,07	,30	-,22	7,17E-02
2	1	7,84E-02	,07	,30	-7,17E-02	,22

(1) score obtained in the 1ª evaluation, relative to the attitude of the participants
(2) score obtained after 6 months, relative to the attitude of the participants

No differences were found between the results obtained in the 1ª evaluation and the results obtained after 6 months (p>0.05) in the variable attitude.

Profile Plots



ANCOVA VARIABLE ATTITUDE
VARIABLE DEPENDENT: SIXACTTIT
FIXED FACTOR: GROUP
COVARIABLE: INIATTIT

Table 2.2.1.11 ANCOVA results of the SIXSCORE VARIABLE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	16,52[b]	2	8,26	16,83	,00
Intercept	24,56	1	24,56	50,03	,00
INISCORE	12,67	1	12,67	25,81	,00
GROUP	6,46	1	6,46	13,17	,00
Error	57,92	118	,49		
Total	7226,00	121			
Corrected Total	74,45	120			

Conclusions: Attitudes

Within the comparisons carried out by the anovas in the attitude variable we find that:

No significant differences were found:
- Between the results obtained in the control group and the results obtained in the experimental group.
- Between the results obtained in the 1ª evaluation and the results obtained after 6 months.

2.2 Evaluation of results: Road Safety Questionnaire

2.2.2 Knowledge

To collect the data several questionnaires were created with a view to collecting as much information as possible on knowledge of the participants. These questionnaires were studied and underwent reliability and validity tests, obtaining acceptable results (α de cronbach=0.62, see annex, separate document, available on request).

The scores in the knowledge variable were obtained by adding the scores of each of the subjects in the questions focusing on knowledge in the questionnaire. The scores range between 0 and 10, 0 being the lowest score and 10 being the maximum. A high score in the knowledge variable would indicate a sound level of knowledge on different aspects related to driving, such as:

- Components and systems of the automobile
- Accident rate data in Spain.
- Traffic regulations
- Ecological driving
- Active safety features
- Defensive driving

Next, the following descriptive data are presented (media and standard deviation) for the knowledge variable, based on an analysis of groups.

Table 2.2.2.1 Variable KNOWLEDGE * GROUP

GROUP		INIKNOWL	WKKNOWL	SIXKNOWL
Experimental Group	Mean	4,56	6,16	5,36
	N	74	77	75
	Std. Deviation	1,45	1,55	1,57
Control Group	Mean	4,51		4,23
	N	72		74
	Std. Deviation	1,44		1,36
Total	Mean	4,54	6,16	4,80
	N	146	77	149
	Std. Deviation	1,44	1,55	1,57

INIKNOWL: 1 evaluation score variable knowledge
WKKNOWL: After the course score variable knowledge
SIXKNOWL: After 6 months score variable knowledge

At a descriptive level, we can observe differences within the experimental group (within groups) and differences between the experimental group and the control group (between groups). Within the experimental group, we can observe that the lowest scores coincide with the first evaluation (pre-test) and the highest scores coincide with the evaluations realized after the training course take place (after one week and after 6 months).

Within the control group, we observe that there are hardly any differences between the results obtained in the pre-test and in the post-test.

As a result of the data obtained, we can conclude that there is a short-term positive effect of the training course and, indeed, a long term lasting effect too because the scores remain high. The results also allow us to conclude that the effects were due to the training course and not to a

learning effect outside of the course when just driving, since no improvement during the 6 month period was noted in the case of the control group.

Table 2.2.2.2 Analysis by gender KNOWLEDGE variable

GENDER		INIKNOWL	WKKNOWL	SIXKNOWL
men	Mean	4,77	6,09	4,90
	Std Deviation	1,45	1,53	1,54
women	Mean	4,10	6,29	4,61
	Std Deviation	1,33	1,60	1,63
total	Mean	4,54	6,16	4,80
	Std Deviation	1,44	1,55	1,57

Table 2.2.2.3 Analysis by gender * group KNOWLEDGE variable

GROUP	GENDER	INIKNOWL	WKKNOWL	SIXKNOWL
Experimental group	Men			
	Mean	4,89	6,09	5,32
	Std Deviation	1,42	1,53	1,64
	women			
	Mean	3,99	6,29	5,44
	Std Deviation	1,34	1,60	1,45
	Total			
	Mean	4,56	6,16	5,36
	Std Deviation	1,45	1,55	1,57
Control group	Men			
	Mean	4,65		4,46
	Std Deviation	1,48		1,31
	Women			
	Mean	4,23		3,82
	Std Deviation	1,35		1,38
	Total			
	Mean	4,51		4,23
	Std Deviation	1,44		1,36
Total	Men			
	Mean	4,77	6,09	4,90
	Std Deviation	1,45	1,53	1,54
	women			
	Mean	4,10	6,29	4,61
	Std Deviation	1,33	1,60	1,63
	Total	4,54	6,16	4,80
	Mean			
	Std Deviation	1,44	1,55	1,57

As we can observe in the results, not much difference exists between the mean scores of men and women. The scores are quite high amongst males. The highest scores coincide with the evaluations after the training (after one week).

Analysis of variance of the variable KNOWLEDGE

ANOVA
KNOWLEDGE

Table 2.2.2.4 ANOVA results of the change variable (knowledge)

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
CHANGE	4,30	1	4,30	2,52	,11
CHANGE * GROUP	17,00	1	17,00	9,96	,00
Error(CHANGE)	237,18	139	1,70		

The change variable= differences between iniknowl variable and sixknowl variable

Table 2.2.2.5 ANOVA Comparisons between experimental group and control group

(I) GROUP	(J) GROUP	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Experimental	Group	,53*	,19	,00	,15	,91
Control Group		-,53*	,19	,00	-,91	-,15

Differences were found between the results obtained in the control group and the results obtained in the experimental group ($p<0.05$) in the knowledge variable.

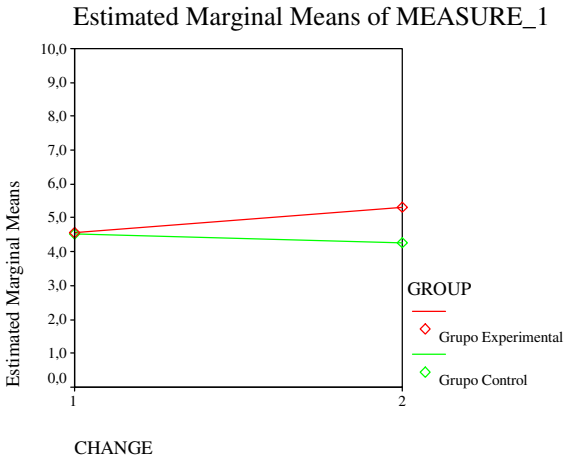
Table 2.2.2.6 ANOVA Comparisons between iniknowl variable (1) and sixknowl variable (2)

(I) CHANGE	(J) CHANGE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-,24	,15	,11	-,55	6,05E-02
2	1	,24	,15	,11	-6,05E-02	,55

(1) score obtained in the 1ª evaluation, relative to the knowledge of the participants
(2) score obtained after the course, relative to the knowledge of the participants

Differences were found between the results obtained in the 1ª evaluation and the results obtained after 6 months ($p<0.05$) in the knowledge variable.

Profile Plots



ANALYSIS OF THE EXPERIMENTAL GROUP (3 MEASURES)

Table 2.2.2.7 ANOVA results only experimental group (knowledge)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
CHANGE	94,18	2	47,09	23,09	,00
Error(CHANGE)	289,56	142	2,03		

The change variable= differences between iniknowlt variable(1), wkknowlt variable(2) and sixknowl variable(3)

Table 2.2.2.8 ANOVA Comparisons between iniknowl variable(1), wkknowl variable(2) and sixknowl variable(3)

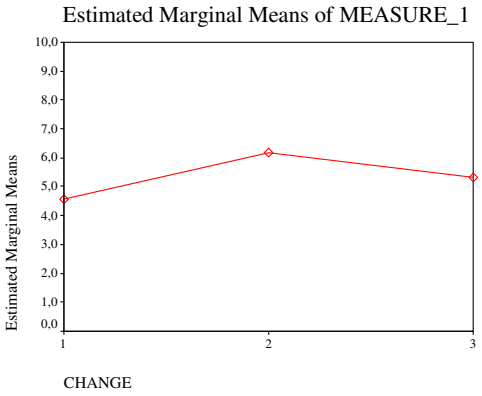
(I) CHANGE	(J) CHANGE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-1,61*	,25	,00	-2,13	-1,09
	3	-,73*	,23	,00	-1,19	-,28
2	3	,87*	,22	,00	,43	1,32
	1	1,61*	,25	,00	1,09	2,13
3	2	-,87*	,22	,00	-1,32	-,43
	1	,73*	,23	,00	,28	1,19

(1) score obtained in the 1ª evaluation, relative to the knowledge of the participants

(2) score obtained after the course, relative to the knowledge of the participants
(3) score obtained after 6 months, relative to the knowledge of the participants

Differences were found between results obtained in the 1ª evaluation and the results obtained after the course and the results obtained after 6 months ($p<0.05$), analysing the experimental group.
We have found differences between the results obtained after a week of carrying out the course and the results obtained after 6 months ($p<0.05$)

Profile Plots



ANALYSIS OF THE CONTROL GROUP (2 MEASURES)

Table 2.2.2.9 ANOVA results only control group (knowledge)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
CHANGE	2,05	1	2,05	1,36	,247
Error(CHANGE)	102,37	68	1,50		

The change variable= differences between inknowl variabl(1) and sixknowlt variable(2)

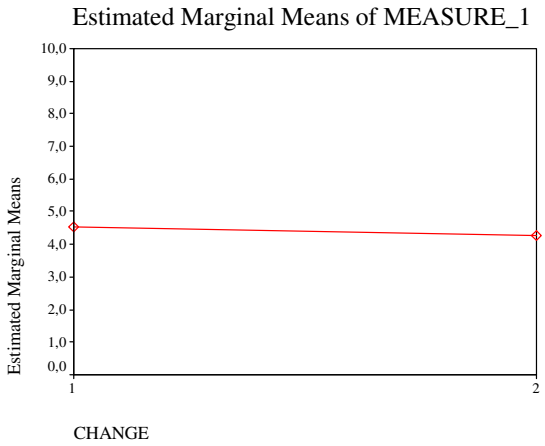
Table 2.2.2.10 ANOVA Comparisons between iniknowl variable(1) and sixknowl variable(2)

(I) CHANGE	(J) CHANGE	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	,24	,20	,24	-,17	,66
2	1	-,24	,20	,24	-,66	,17

(1)score obtained in the 1ª evaluation, relative to the knowledge of the participants
(2) score obtained after 6 months, relative to the knowledge of the participants

Differences were found between results obtained in the 1ª evaluation and the results obtained after 6 months ($p>0.05$), analysing only the control group.

Profile Plots



ANCOVA VARIABLE KNOWLEDGE
VARIABLE DEPENDENT: KNOWLEDGE (SIXKNOWL)
FIXED FACTOR: GROUP
COVARIABLE: INIKNOWL

Table 2.2.2.11 ANCOVA results of the SIXKNOWL VARIABLE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	50,29[b]	2	25,14	12,13	,00
Intercept	194,94	1	194,94	94,04	,00
INISCORE	13,04	1	13,04	6,29	,00
GROUP	36,55	1	36,55	17,63	,00
Error	286,04	138	2,07		
Total	3580,33	141			
Corrected Total	336,33	140			

Conclusions: Knowledge

Within the comparisons carried out by the anovas in the knowledge variable (see annex 12 for questions related to knowledge) we find that:
Significant differences exist:
- Between the results obtained in the control group and the results obtained in the experimental group.
- Within the experimental group. Significant differences exist between the results of the 1ª evaluation and the results obtained (a week) after the course and after 6 months.
- Within the experimental group, between the results obtained a week after carrying out the course and the results obtained after 6 months.

Significant differences do not exist between the results of the control group in the initial evaluation and the control group after 6 months.

3. Conclusions

Objective of evaluation:
The aim of the evaluation was to determine whether this post-licence driver training course for novice drivers was effective in terms of improving knowledge, skills and a positive change in attitudes towards traffic and road safety. The results obtained in the study show:

- ✓ **There were significant differences in the skills and behaviours** of the participants as a result of the training course, so we can conclude that the course had a positive short term effect. In the same way, this change is maintained for a longer time 6 months period, so we can also conclude that there is a lasting effect of the course.

- ✓ **No significant differences in attitudes were found** amongst the participants in the training course, so the results are not so favourable as expected and no major conclusions can be drawn in this respect.

- ✓ **There was a significant difference in knowledge** amongst the participants in the training course, so we can conclude that the course had a positive short term effect. In the same way, this change is maintained for a longer time 6 months period, so we can also conclude that there is a lasting effect of the course, although the effects have decreased slightly over this period.

The results obtained confirm the initial hypothesis of the study, namely that this driving course is interesting for novice drivers and that it helps to improve their road safety in several ways.

It is often claimed that the level of knowledge amongst novice drivers is insufficient to drive safely. With this study, we can conclude that with such a course it is possible to improve the level of knowledge of novice drivers.

It has also been claimed that novice drivers' skills need to be improved. With this study, we can conclude that with such a course it is possible to improve the level of skills of novice drivers.

In the same way, it has been said that attitudes in traffic are especially relevant to road safety. With this study, we cannot conclude that with such a course it is possible to improve attitudes of novice drivers and further studies should be done.

In summary, this study concluded that this driving course is effective for novice drivers.

In the short term, we can say that this course is positive for those drivers who take part in them and, over a longer period, that the courses add a social benefit to the society and road safety in general.

In terms of improving the current study, the following points may be noted:
- More time and economical resources would have allowed for a better study.
- It would be useful to correlate the results with the ones obtained in other countries
- If a new study were to be made, the initial sample would have to be increased in order to avoid statistical mortality (a high dropout rate).

As a final comment, it is recommended to focus on the training of the trainers, and to recommend the presence of a psychologist specialised in driver psychology.

Reliability and validity

To collect the data, several questionnaires were created with a view to collecting as much information as possible on knowledge, skills, behaviour and attitudes of the participants. These questionnaires were studied and underwent reliability and validity tests, getting acceptable and good results.

On-road evaluation form α=0.95
Evaluation Questionnaire α=0.62 (knowledge) and α=0.59 (attitudes)

The evaluation questionnaire was divided into two parts. One evaluated theoretical knowledge, and the other one attitudes.

The best results of the reliability test are obtained for the on-road evaluation form. We consider, according to the results obtained, that these tests are demanding and reliable enough to be able to draw conclusions in our study.

Referring to the attitude variable, we should take into account the difficulty of designing a questionnaire for attitude measurement, although the reliability analysis for that questionnaire obtained an acceptable result. Despite this, the results of the attitude variable are not as favourable as expected.

The size of the participant sample in the study was designed to be big enough to ensure the representativeness of the group in relation to the general population of novice drivers. The selection criteria were totally unknown to the participants to avoid the possibility of a selection bias. In the same way, there was a risk that the subjects participating in the project had a positive attitude to road safety and good behaviour in traffic, and were thus not representative of the whole population. For this reason, participation in the Road Safety training programme was not an incentive in itself. The only incentives for participants were based on other interesting aspects for young people (a lottery to win a Renault Clio, some gifts, etc.)

7. SUMMARY AND ANALYSIS OF EACH TRAINING PROGRAMME

Due to the length and complexity of the individual country reports in the preceding chapter, this chapter aims to provide a summary overview of each programme, in addition to an analysis of the training programmes including best practice examples selected by the project manager. The analysis component relates only to the training itself, not the evaluation (see next chapter).

7.1 AUSTRIA: Executive Summary

Participants
Due to the fact that the multiphase system for novice drivers has been obligatory in Austria since 1st January 2003, a deliberate selection of participants was not necessary. Therefore the group samples for this project were chosen randomly.

Training programme
The second phase training in Austria consists of the following modules:

- Two on-road feedback drives (before and after the track training)
- A track training (on a closed track)
- A psychological group discussion

Trainers
The on-road feedback drives for novice drivers is accompanied by driving teachers, the track training is led by instructors and the group discussion is conducted by psychologists.
All involved professions have to fulfil several requirements (e.g. education, age, etc.) which are defined by law.

Evaluation design and data collection methods
The evaluation design (see Table 1) is based on three levels: a process evaluation for both trainers and participants with regard to the track training and the group discussion, a wide scale survey concerning driving attitudes, beliefs and other self-reported data and statistical data from a file of the Central Licence Register concerning all novice drivers in Austria. The predominant collection method was the usage of questionnaires.

Table 39: Evaluation design and data collection methods

	evaluation type	data collection	when
1a	Process evaluation: participants	questionnaire for MPE (="Multi-Phase-Educated") participants	before & after track training
1b	Process evaluation: trainers	questionnaire for MPE trainers	after track training
2	Wide scale survey	control group (SE="standard" education) from "BASIC"(a previous EU-project): questionnaire intervention group (MPE): questionnaire	before and after the introduction of the multiphase system
3	CLR data	Central Licence Register:	Cut-off date: 1st of April 2004

Analysing methods
For this evaluation only non-parametric tests were used since basic requirements for parametric test were violated.

Results and conclusions
Novice drivers who completed at least two modules of the multiphase system were generally satisfied with the whole measurement although it is obligatory. This circumstance can be interpreted as evidence for high acceptance the multiphase system in Austria.

For the track training day, most participants mainly expect to learn to master risky situations better. Also the practical part of the track training day was assessed as most applicable for every day driving. Furthermore, the results show a different view on the importance of several skills between instructors and participants: For example, the ability to correct a skidding car was rated significantly more relevant for real traffic for novice drivers than for instructors, although all skills were considered as very important for safe driving. Nevertheless, it can be concluded that participants may have received a counterproductive message concerning traffic safety during the track training, i.e. that safe driving is based on manoeuvring skills rather than on an anticipatory driving style.

The results of the wide scale survey show that the reduction of practical and theoretical hours of the standard education didn't have statistical significant influence on the pass-rates (number of attempts) of the driving exam, neither on the theoretical test nor on the practical test.

No big differences were found between standard-educated and multiphase-educated novice drivers concerning self assessment of driving style and driving behaviour, offences or accidents. The only differences occurred regarding female persons: they described themselves as more careful drivers and reported less speeding offences.

7.2 FRANCE: Executive Summary

The participants
396 young members of MACIF insurance company, aged between 18 and 2 years old and having between 4-6 months driving experience, participated in the NovEV project. These young drivers were split into 3 groups : 124 in the experimental group, 87 in the control group and 124 in control group 2. Control group 2 was unaware that it was being monitored, whereas the other two groups had expressed an interest in participating actively in a road safety training programme.

The training
Experience gained in the past by ECF suggested that the programme should be spread over 2 days. These two training days were separated by a 4 month interval. This allowed for more intensive debates and exchanges between the participants.
The training programme contained information, and discussion on different risks (either subjective or objective). It alternated between workshops, on-road sessions and track-based modules. The programme takes into account the hierarchical model of driving behaviour and is particularly focused on levels 3 and 4 of the GDE (goals for driver education) matrix.

The trainers
The whole programme depended heavily on the quality of the discussion and on the pedagogical quality of the training. The 5 trainers used were road safety professionals who were qualified and experienced in giving training to groups of young drivers. They trained in pairs during the

entire programme. In order to help them and to retain a coherent approach amongst the different trainers, a trainers' guide was developed especially for this programme.

Feedback on the training
The organisers, trainers and participants all rated the experience positively.

The evaluations
The main objective was to measure and to compare changes in skills, attitudes, knowledge and driving behaviour amongst the participants who actually took the training, and those who did not.
The participants were monitored over a period of 11 months using specially designed questionnaires, as follows :

- Pre-training questionnaire (experimental and control group 1)
- Post-training questionnaire (experimental and control group 1)
- MACIF accident monitoring (for the 3 groups)

Results
Positive changes in the experimental group :
Significant positive change in awareness of risks linked to driving habits (MALES)
Significant positive change in driving skills for defensive driving (MALES)
(Slight) trend towards less frequent risky driving situations (MALES)

Stability of control group.

Conclusion
We can reasonably conclude that the development of the two groups shows an increase in risk awareness in the experimental group. This helps to delay the phenomenon of overconfidence which is so often observed amongst novice drivers.
Otherwise, the control group, which was followed statistically but not involved in the training, remained stable in its results, despite its clear investment in road safety (by wanting to take part)

7.3 GERMANY: Executive Summary

In April 2003, a voluntary second-phase training programme for probationary (novice) drivers (FSF) was introduced in Germany by law as a pilot project. Between 2003 and 2010, the FSF project aims to evaluate how, if at all, it contributes to reducing novice drivers' accident risk. To date, 13 out of 16 federal states in Germany have joined the pilot project and have started the training which offers an incentive in the form of a one-year reduction of the probationary period for the novice drivers who participate. The FSF courses actually started in spring / early summer 2004.

Before then, some preparatory work had already been carried out by the DVR (Deutscher Verkehrssicherheitsrat = German Road Safety Council). The DVR developed the manuals and subsequently trained the trainers to coach the seminar leaders (driving instructors). When the training was introduced in practice, approximately 1,500 seminar leaders and 200 track instructors had already been trained to implement the FSF model. In the first five months of training, about 200 novice drivers took part in the FSF-courses. The training is composed of five sessions, including three group discussion sessions, one 'training and observation' drive on public roads, and a track-based training programme. Overall, the demand for the courses within the target group of young drivers has been rather low so far.

BASt collected the addresses of all people involved in designing and implementing the FSF programme so far. Six quasi-identical questionnaires were developed for the six groups and sent

to the persons involved: programme authors, the trainers of the trainer, the seminar leaders and track instructors, and the participants (novice drivers). Data collection took place in June and July 2004. Due to the tight deadline, no pre-testing or follow-up-testing could be performed. The main focus of the evaluation was on the perception of the training itself among the different categories of people involved and on to what extent the programme is transferred to the participants in such a way the programme authors intended it.

The German evaluation project was a process evaluation with a single measurement. Different views on the programme with regard to the organisation, implementation and achievements were collected from the six different groups involved, such as the course designers and the participants of the programme. The effectiveness of the programme for the young drivers was also analysed by comparing the learning goals set down in the manual with reported self-assessment and an evaluation of the participants' success in implementing the goals of the training in practice.

The results of the study provide an indication of how accurately the FSF training concept was implemented in practice. The results show that most of the programme is being performed as the authors intended. The task reports and ratings of the importance of the programme modules correspond closely to the authors' concept and manual, the implementation as carried out by the trainer, and the participants' experience with FSF. Similar statements on the FSF modules were found in the six groups surveyed. The training was generally rated positively by the participants. The participants claimed to have developed and used several intellectual and behavioural strategies for safe driving, which is the main aim of FSF.

Attention should be turned to the fact, that participants reported an unwanted improvement of their abilities to master difficult traffic situations. Also the training of the track instructors should be revised because they perceived and implemented the track training course with other goals and intentions than was specified by the authors.
Further efforts to motivate novice drivers to participate in FSF should be made. Suggestions on how to do this are made in the conclusions.

7.4 NETHERLANDS: Executive Summary

Participants
After an appeal by mail and telephone, 376 young novice drivers agreed to participate in the project. Unfortunately, during the course of the project, many of the participants dropped out. Out of 376 young drivers that initially agreed to participate, only 127 (33%) completed all parts of the project.

The participants who did not want to participate, those who dropped out, and those who finished all parts of the project were compared for a number of variables. This led to the conclusion that there was no major problem with selective drop-out. Naturally, the groups did differ on at least one aspect, namely for one reason or another some completed the project and others did not.

Training programme + objectives
The second phase training consisted of the following modules:

- *An on-road feedback drive*
  The objective of the feedback drive was to present the driver with feedback about his driving performance. It was different from instruction drives, as the instructor confronted the driver with his "expert" observations in order to make the participant "think" and reflect. So he did not tell the participant what to do, but encouraged him to draw his own

conclusions. During the first feedback drive the participant and instructor were accompanied by a second participant who rode along as a passenger. The drive was followed by a discussion between instructor, passenger and driver.

- *Training on a closed track*
  The objective of the track training was for participants to experience the limits of their skills in vehicle control and to share these experiences with other group members.

- *A group discussion*
  The objective of the group discussion was to stimulate recognition of potentially hazardous situations in rather "normal" social situations. The discussion was based on video sketches, depicting typical situations (incidents rather than accidents) involving young drivers (men and women). The moderator encouraged the youngsters to reflect on the events.

- *An evaluation on-road feedback drive (about a month later)*
  The objective of this second feedback drive was the same as the first feedback drive, that is to present the driver with feedback about his driving performance.

Evaluation design and data collection methods

The effect of the track training and group discussion was studied using a before-and-after design with a control group. Participants were randomly assigned to the control or the experimental (treatment) group. The *control group* participated in both feedback drives. In addition to the feedback drives, the *experimental group* also participated in track training and in a group discussion.

Evaluation design and data collection methods

	Training programme		**Instruments**
	Experimental	*Control*	
December 2003 **Pre-test** *One month before training*	Questionnaire	Questionnaire	**Questionnaire** Contained items on risk awareness, self-assessment of skill, and situation judgements
January 2004 **Training day**	Pre-test feedback drive	Pre-test feedback drive	**On-road observation form** An assessment tool to describe the driving performance of a driver. The driver himself and the driving instructor completed these forms after the feedback drive. **Driving Assessment** Assessment by the instructor of the quality of driving in three fields: vehicle control, driving skills, and calibration skills
	Track Exercises		
	Group discussion		
	Questionnaire	Questionnaire	**Questionnaire** Contained items on risk awareness, self-assessment of skill, and situation judgements

		On-road observation form An assessment tool to describe the driving performance of a driver. The driver himself and the driving instructor completed these forms after the feedback drive.
Post-test feedback drive	Post-test feedback drive	**Driving Assessment** Assessment by the instructor of the quality of driving in three fields: vehicle control, driving skills, and calibration skills
		Satisfaction questionnaire This questionnaire contained questions on how satisfied participants were about the different components of the training day and the feedback drives.

Results & Conclusions by instrument

Satisfaction questionnaire
Young drivers were not motivated to participate on a voluntary basis in a second phase training. However, once in the course, novice drivers were enthusiastic about the training day. Within the training day, the group discussion was rated as the least attractive part, while the feedback drive was about as attractive and useful as the track training. The message of the second-phase training was well-understood. There were no indications that the young, novice drivers overestimated their skills, as a result of the training.

Questionnaire
The questionnaire contained items on risk awareness, self-assessment of skill and judgements of traffic situations on photo. The results from the questionnaire are somewhat unclear; some effects of the training were found, but not consistent and not always in the expected direction.

In line with expectations, the items concerning risk awareness confirmed that young drivers do not seem particularly concerned in general, and especially not about driving too fast. A least 60% of the respondents are not concerned about driving too fast. On the other hand, it turned out that young drivers are, overall, rather confident about their driving skills. At least 30% of the participants believe they are (very) strong in all skills, and in some skills more than 60% believe they are (very) strong.

It was expected that these opinions would improve as a result of the training day. Detailed analyses showed no effect of training on these variables. Further research is needed to demonstrate that the questionnaire itself is sensitive enough to register changes as a result of a short term intervention. The fact that there were significant gender differences in these issues, led to the conclusion that this part of the questionnaire possibly measures more stable attitudes or personality traits (which could not be changed with a one-day training course or within the period of a month).

On-road observation form
After the feedback drive, an on-road observation form was filled out by both the instructor and the participant, which contained items on driving skill and assessment of complexity of the driving task. The young drivers' assessment of their own driving skills and task complexity did not change as a result of training. This implies, that the objective of the course to inform young drivers about their limited skills and the high complexity of the traffic situation did not result in a more cautious self-estimation. On the positive side, this result indicates that the training day and more in particular the track training did not lead to a higher estimation of skills and a lower estimation of the complexity of the driving task.

To study the accuracy of the driver's self image, their self-estimation scores were compared with the instructor's assessment of the young driver's competencies. On "vehicle control and general skills", instructors and participants did not differ in their assessment neither on the pre-test nor on the post-test. As expected on "safe and defensive driving" in the pre-test, participants rated their performance higher than the instructor did. As the course was directed at improving self-assessment skills, it was expected accuracy to improve in the sense that their assessment would be more in line with that of the instructor after the training. This was not the case.

Generally, from the results from the on-road observation form, it can be concluded that while the instructors did see some improvement as a result of the training, the participants did not.

Driving Assessment
Task conditions between control group and experimental group differed systematically on the pre-test. Therefore, it cannot be excluded that the observed difference in task performance between control group and experimental group is a reflection of these test conditions rather than a significant difference between the two groups.

Within the experimental group, the performance of the participants of the two different training locations differed significantly. This, despite the fact that at both locations the participants had received exactly the same training (on paper). Where performance at location A was improved by training, driving performance at location B got even worse. Because the test conditions for the participants of the two locations were the same, this result is reliable.

The process evaluation indicated that despite their organisation's involvement in the NovEV project, the trainers from location B did not share the same opinion on the definitions of a "useful" training. As a result, these trainers had to give a type of training they did not believe in. This could have (subconsciously) affected the way they gave the training, or the way the participants perceived the training. Research has shown (ADVANCED, 2002) that any education, looses its strength if the educator is not absolutely convinced about what he/she is teaching. Moreover, that the effectiveness of the education is largely dependent on the person, the beliefs of the teacher, and his behaviour (Hale and Glendon, 1987). For a more detailed discussion of the role of the "teacher", see the ADVANCED report.

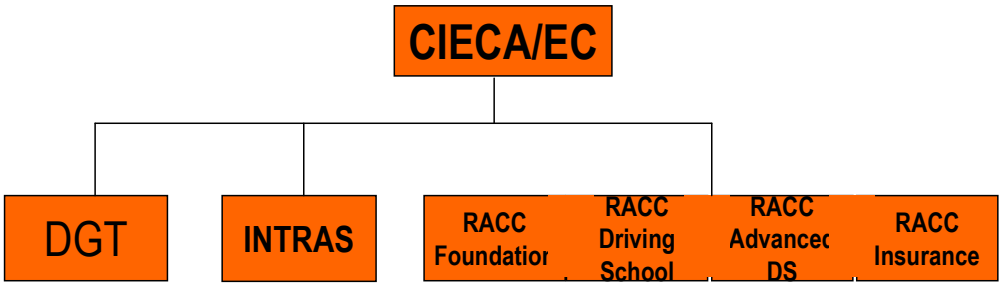
General conclusions
In the Dutch pilot, the recommendations of the ADVANCED report were closely followed with respect to the content of the course and the evaluation of its effects. However, as stated earlier, in practice these recommendations were not always followed in one of the two locations.

In this study, it has been demonstrated that, on the one hand, the second phase is recognized by the participants as a useful and necessary part of their driving career. On the other hand, the high refusal rate demonstrates that youngsters are not interested in participating on a voluntary basis. The effects of the course are limited, and can even be negative, if trainers are not fully equipped to give the course, indicating that a much greater effort is needed in training second phase trainers than has been the case in this project.

7.5 SPAIN RACC: Executive Summary

The NovEV pilot trial of the RACC Automobile Club took place during the period between January 2003 and May 2004, and involved 621 participants from three provinces of Spain: Madrid, Valencia and Barcelona. The aim of the project was to evaluate a post-license training course in order to assess if it can positively influence the behaviour of novice drivers. An evaluation strategy was planned, based on a experimental research design with experimental and control groups assessed at two points, before and after the training, with regard to a number of variables related to safe driving.

The partner structure that RACC built in order to develop and implement the NovEV pilot trial was led by the European supervisor, CIECA, whereas at a national level, the traffic authority DGT (Dirección General de Tráfico) supported the pilot trial and will use the results for the future development of post-license courses. The training design and its implementation at national level was managed by the RACC Automobile Club and INTRAS (University of Valencia).



NovEV pilot trial: structure of partners

In January 2003 a massive marketing campaign by post was addressed to more than a thousand young drivers who were all policyholders with HDI car insurance company. These potential participants fulfilled the following specifications: aged 18 to 24, less than 3 years driving experience and living in Valencia, Barcelona or Madrid provinces. The letter informed them about a pilot trial in which they were invited to participate and explained what it involved (at least two tests over a one year period) and what would they get (possibility to take a training course, and to win a car in a lottery).

Phone calls followed the marketing campaign in order to recruit participants and to conduct a short interview-questionnaire (see 1.7, selection questionnaire) that would provide the basic background from each participant. The information taken from the phone questionnaire supplemented the information from the insurance company database which provided the basic variables needed to segment the sample into two balanced groups in terms of age, gender, educational background, driving experience and vehicle use.

At this stage two groups were formed: a control group and a test group that totalled 621 people. A pre-test (see 1.7; pre-test) of driving behaviour was sent to them and 350 answered within the deadline (183 from the test group and 167 from the control group). Due to an unexpected high rate of dropouts, the participation in the project was extended not only to HDI insurance holders but also to other members of the public, mainly recruited at driving schools in Valencia and Madrid.

The training days took place in Barcelona (3 days), Valencia (3 days) and Madrid (1 day) in July 2003, involving the 183 members of the test group. It consisted of a one day training during which all participants had to take three areas of training: track training, on-road feedback drive

and psychological workshop. Before the training began, the RACC conducted a rehearsal day in order to train the trainers and to improve certain organisational aspects.

The track part consisted of two parts: performing emergency braking with and without ABS on slippery and rough surfaces on one hand, and experiencing an exercise in which participants were distracted by mobile phones and peer pressure. The on-road section combined urban and rural roads on a pre-defined circuit in which each participant had to drive for 20 minutes. In the workshop section, the most important sociological and psychological aspects that affect young drivers were introduced and discussed. Every session except the workshop had a feedback session during which the trainees had the opportunity to interact with the trainers and to have their questions answered.

According to the project guidelines, a period of 5 months then elapsed during which no contact was had between RACC and the participants, in order to allow for consolidation of any attitudinal improvements as a result of the course among the members of the test group.

During the period from December 2003 to end of January 2004 the post-test (see 1.7; post-test) was sent to the 350 people still involved in the NovEV pilot trial. The final participation figures after dropouts from the pre-test and the post-test was 263 novice drivers, namely 126 from the test group and 137 from the control group, of which 66% were from the HDI insurance database and 33% were from driving schools.

During the period from January 2003 to March 2004, any reported accidents were monitored amongst participants from the HDI insurance database. Despite being an unstatistically consistent result, the query showed that 4 participants from the control group were responsible for an accident, whereas only 1 participant from the test group was responsible for an accident.

Two basic methods have been used to collect the data from participants: phone interviews (recruiting questionnaire) and post (pretest and post-test). One last source of information has been the database from HDI.

After all the data was collected, a comprehensive statistical analysis was carried out on the data from the final sample of participants that completed the two driving behaviour tests. The original data provided by the HDI database and the recruiting questionnaire were also used to perform the analysis by providing segmentation variables and to detect any self-selection bias.

The evaluation strategy was based on an experimental research design with experimental and control groups assessed at two points, before and after the training, according to a number of variables related to safe driving. An univariate ANOVA model was used to analyse the data of our mixed between-within design for each one of the five scales considered.

Data analysis results showed statistically significant differences between the control and test groups for the “Skills for Careful Driving” scale, meaning that the mean score in this scale was higher for the test group than for the control group after the training. This result goes in the expected direction given that, as reported in earlier studies, self-evaluation of skills for careful driving is inversely related to accidents. Positive differences between the test and control groups were also found for the other four driving behaviour scales, but these differences did not appear to be statistically significant, so they could have occurred by chance.

Finally, data analysis of the course feedback obtained from the participants of the test group showed a rather positive evaluation of the course and the course results. The first conclusion is supported by mean scores over 4 (in a 1 to 5 scale) for the items related to the course organisation, contents and tuition; the latter through the mean scores for the improvements which were reported by participants, which were significantly higher for the items related to self-awareness about risks and bad driving habits than for driving techniques and skills.

7.6 SPAIN RACE: Executive Summary

The general target of this study was to evaluate scientifically the influence of second phase training courses on novice drivers (once the driving licence has been obtained and some driving experience accrued) and to measure any changes related to skill, knowledge, behaviour and attitude as a result of the course.

The sample participating in the study was composed of 154 subjects. This sample was selected according to the following selection criteria: category B licence holders for between one/two years and a minimum of 5.000 km of driving experience. Once selected, the participants were randomly assigned to two groups:

- Experimental group (77 participants)
- Control group (77 participants)

The experimental group took part in the training. The course contents were focused on a few very clear messages, especially oriented to road safety. The Training Programme was composed of three different modules: class (theoretical contents, discussion), track (practical contents), real traffic (assessed driving). The control group did not participate in the training. The aim of the control group was to establish the base line in order to determine which part of the change achieved in the experimental group was due to the training and which part was due to the driver's natural development.

Evaluations were conducted at 3 stages with a view to establishing the differences between the two groups over a period of six months. The first evaluation (pretest) took place before the training programme in order to establish a base line and to be able to compare later evaluations. Two further evaluations followed after the training programme: after a week (to evaluate any results over the short-term) and after six months (to evaluate results over the medium to long term).

The following data collection methods were used for this purpose: a road safety questionnaire and an evaluation in real traffic (driving assessment on public roads) in order to compile as much data as possible related to current knowledge, skills, behaviour and attitudes of the participants. The data analysis methods used in the study were a descriptive analysis and ANOVA.

In the on-road evaluation, a significant improvement in general driving skills within the experimental group was found as a result of the training. In the attitude variable, no differences were found. According to the questionnaire feedback, there was a significant improvement in knowledge within the experimental group as a result of the training.

CONCLUSIONS:

Analysing the results obtained, we can conclude that there was a significantly higher change in knowledge, skills and behaviour (in the attitude variable the results are not so favourable as expected, so no conclusions can be drawn from these) in the experimental (training) group than in the control group. Therefore, the training was seen to have a considerable effect on participants in the short and medium term (6 months). We can thus conclude that there was a positive effect of the course on novice drivers.

In the skills and behaviour variables we found that:
The training improved the participants' driving skills and behaviour in comparison to the control group

In the attitude variable we found that:
No differences were found between the results obtained in the experimental group and the control group

In the knowledge variable we found that:
The training improved the road safety knowledge of the participants in comparison to the control group.

7.7 Analysis of individual training programmes

With the benefit of hindsight, a number of important lessons have been learned with regard to the design and implementation of 2nd phase training. The following passages present the strengths and weaknesses of each of the schemes evaluated during the NovEV project. The vast majority of these observations come from the scheme managers themselves; others have been noted by CIECA and / or the NovEV independent evaluation advisor. Ideally, the training programmes could have been analysed in the context of the GDE matrix in order to see how far the Advanced guidelines were being respected. In practice, however, this is a complex and subjective process (and the differences between the programmes on paper were small) so it was decided not to do this.

7.7.1 Austria

The new, post-licence component of the Austrian multiphase is based on the Finnish approach but is supplemented with an additional feedback drive after the track training/ group discussion, as well as before. This design, in addition to the intervention period (all 3 modules within 1st year of driving licence) is very positive. Entering into more detail, however, the balance between the different modules, particularly with regard to the psychological group discussion and the track training, appears uneven. Only 2 hours is allocated to the group discussion, compared to 6 hours for the track training (1 hour theory and 5 hours practice). Furthermore, the results show that many participants claim that a major focus of the training was on mastery of vehicle control – this is not an objective of the training. The feedback drive has been difficult to implement properly, due to the fact that by the time the multiphase was introduced, no proper training had been given to driving instructors on how to carry out this module, and the driving assessment form to be used was still in the design phase. Initial feedback indicates that driving instructors are operating in the feedback drives much as they would during normal, pre-licence lessons (i.e. instructing rather than coaching). Based on its design and content, the group discussion should be praised due to its focus on participant-centred methods and attention to the higher levels of driver behaviour.

In terms of the overall management of the Austrian multiphase, two weaknesses can be observed. Firstly, there is still no handbook to ensure the quality and coherence of the training given by a panoply of different training organisations. Individual training organisations may have their own handbooks but these are likely to differ considerably in practice. This means that the training itself is probably different across the country. Secondly, the quality assurance committee includes the organisations that are being controlled (namely OEAMTC and ARBO). CIECA considers it important that the quality assurance of a 2nd phase programme be independent from the organisations directly involved in the training.

7.7.2 France

As one of the pilot projects in NovEV, France/ECF is to be commended for adopting a full 2-day training programme with a 4 month interval between. This structure allows not only for a longer period of training and instruction but also an extended support structure for novice drivers following the licence. ECF benefited from the considerable experience and skills, particularly with regard to coaching, of trainers who have been trained first and foremost as ‘animateurs’ or group leaders for young people, in addition to holding driving instructor qualifications. France was also the only country where trainers worked in pairs during the group discussion and track training sessions, and to good effect. The majority of the themes discussed in the group discussion seemed sound, despite some initial reservations expressed by CIECA on

specific issues (presentation of road safety policy, session on visual illusions…). The track training was well designed and implemented, there was no trace of manoeuvring skills training and the discussion amongst participants about each other's vehicles was innovative. The feedback drive was not seen in practice: 6 passengers and a trainer drove on open roads in a people carrier. ECF felt that the immediate comments given by the passengers to the driver allowed for a precise evaluation of the driver's driving style.

Two small but important observations for the future were made during the training. The first is that the audit on the second training day should be replaced with a feedback session on the experiences accrued by the participants during the intervening period between the training days. The second is that, in order to optimalise the conditions for learning during the track training, some form of shelter is necessary next to the track (poor or cold weather, too hot weather, etc).

7.7.3 Germany

A considerable amount of thought and research, resulting from a number of years experience with post-licence driver training, has led to the creation of a very sound voluntary 2nd phase programme in Germany. Positive aspects include strictly defined manuals for trainers, high and relevant trainer qualifications, considerable ground covered in the 5 modules, and more attention to level 3 and 4 (GDE matrix) issues than in other NovEV schemes. The course is also very interactive (including role plays) and is the only course in NovEV to offer the opportunity to participants to request specific training on issues of personal concern. Importantly, the training puts a high emphasis on the development of personal strategies for safe driving.

The course is long and spread out over time. This is both positive and negative. Due to its voluntary nature, the standard 8 week programme may deter young drivers from registering. There is also concern that the incentive to participate (a one year reduction in the probationary period) is attracting drivers who register solely for this purpose. Young traffic offenders (who have already had their probationary period extended by two years) are overrepresented in the voluntary 2nd phase programme at this stage[50], although this may be explained by the publicity made about the FSF in the traffic offenders seminars that these drivers were obliged to attend.

One other potential issue of concern is that track trainers, seminar leaders and, consequently, participants may not have the same understanding of the goals of the track training as the authors. A statistically significant group said they were better at mastering dangerous situations as a result of the whole measure. The improvement of manoeuvring skills was not a major objective of the track training, so this potential for overconfidence amongst participants should be monitored over time and the goals of the practical safety exercises should be communicated more clearer to the trainers and the seminar-leaders in the way the authors wanted. (This incorrect perception of the aims of the track training exists despite a great deal of effort being spent training trainers and establishing detailed trainer handbooks).

7.7.4 Netherlands

Despite the considerable preparatory work on paper, the ultimate outcome of the Dutch Young Drivers' training was mixed. The management 'blueprint' successfully established the framework conditions for the training, and innovative initiatives were taken in the build-up to the training (development of the website, traffic situation diaries, self-assessment as part of the driving assessment and of the video containing sketches of typical novice driver situations). The problem, in one training location in particular, was the translation of the work on paper into

[50] Such drivers account for about 25% of participants in the voluntary programme, compared to around 5% nationwide.

practice, (as was the inclement weather which seemed to discourage many participants from attending the training). A lack of finances meant that only (very) limited training-of-trainers was available, and this shortcoming was sometimes noticeable in practice. For similar reasons, rehearsals were not held, thereby preventing any opportunity for the project management to identify and address weaknesses before the training began. In addition, nobody from the project management systematically attended the training days, in order to monitor the functioning of the programme, give feedback, etc.

The experiences in the other training organisation, where positive results were recorded, were far more successful. There, the objective of the programme was realised – the participants' perception of their risk awareness was more in line with reality as a result of the training (better 'calibration').

Satisfaction ratings were generally high, particularly with the feedback drives. The one exception was in the training location where the negative results were registered. The trainers there were not properly trained according to NovEV / Advanced guidelines and this situation was exacerbated by conflicting messages from the management who were not convinced about the Advanced approach to 2nd phase training.

Considerably more training, monitoring and rehearsing, including better coordination and follow-up between the project managers and the course providers, will be necessary to ensure the successful introduction of a nationwide 2nd phase programme. It may also be useful to note that a tendency to involve a large number of organisations may serve to undermine the coherence and uniformity of the training in practice. The overall experience of the programme in the Netherlands is positive for two reasons: firstly, one training organisation had good results, and, secondly, the problems encountered in the other training organisation are useful lessons for the future.

7.7.5 Spain RACC

In many ways, RACC paid meticulous attention to the Advanced guidelines in practice and developed a tight, well-managed 1-day training programme, which has the advantage of being mobile. The full programme was rehearsed and modified accordingly, before the actual training began. One of the two track experiences was particularly innovative (i.e. the slalom focusing on level 3 issues) and experienced seminar leaders (psychologists) were used in the group discussion module. The presence of young trainers was also positive.

More time could have been allocated to each participant during the feedback drive, and the feedback from the trainer could have been more regular and forthcoming (highlighting both strong and weak points). Improvements in the group discussion are mostly on the micro-level: adding a break, encouraging more interaction between the participants and limiting the number of subjects but entering into more depth. The track trainers and feedback drive trainers would benefit from more coaching training (there was a tendency rather to 'instruct'). Following the training, a decision has been made by RACC to lengthen each module in future trainings to 2 hours apiece, instead of 90 minutes.

If there is one obvious weakness with the RACC training, it is the limited length of the training (1-day only). Whilst CIECA understands the practical structure of the training from a commercial perspective, it would probably be more effective if the intervention period could be lengthened, for example, by adding a self-evaluation questionnaire to be filled out by the participant some time before the training.

7.7.6 Spain RACE

The RACE course was different from the other NovEV schemes in 3 main respects:

1. a disabled (and young) trainer was used from a Spanish NGO raising awareness of spinal and brain injuries (largely due to traffic accidents)
2. the group discussion lasted far longer than the track training (4 and 2 hours respectively)
3. The support period for the young driver was considerable, in that driving audits took place on two separate occasions following the training

The rehearsal revealed that too much information was being presented to the participant, so the RACE training focused on a smaller number of simple messages. A detailed course manual was made available to the participants to encourage on-going learning. The track training lacked momentum due to a lack of cars, but this is more an organisational issue than a substantive one.

As was the case with the French and RACC trainings, the RACE programme benefited from the presence of the same official at each level of the process: the design, management and implementation of the training. This ensured coherence and uniformity in the programme.

7.8 Best practice examples from NovEV programmes

On the basis of the Advanced guidelines, and including the principles of the GDE matrix (in particular the levels 3 and 4 and the self-evaluation column), CIECA has selected examples of best practice from the NovEV training programmes. These examples include not only activities in the training programme itself (content level), but also examples of good practice at a management (organisational) or conceptual (design) level.

7.8.1 Design level

- The Austrian psychological group discussion. See pages 6-8 of the Madrid meeting minutes in annex 4, meeting 2. An English or German version of the trainer's guide is available at CIECA on request.
- The quality control system in the German 2nd phase programme[51]

7.8.2 Organisational level

- The presence of a specially trained coach who is handicapped as a result of a road accident (Spain RACE)
- Two trainers working together (France ECF)

[51] Based on the following sequences:
- Basic research into young drivers problems
- Definition of goals of training
- Development of training programme and training-the-trainer requirements
- Training the trainer
- Implementation of trials and evaluation of training
- Optimization
- Implementation of actual training
- Ongoing quality control (inspections, evaluations)
- Ongoing training of trainers

- The selection of simple 'track' areas for the track modules (Spain RACC and Spain RACE) [52]

7.8.3 Content level

Feedback drive:

- The G-CAM in-built camera and data monitoring system from the Belgian project which allows specific driving situations to be recorded and replayed (with analysis) after the session
- Practising situations already identified as a weakness by individual participants (Germany)
- The self-evaluation part of the driving assessment form used in the Netherlands

Group discussion:

- The video of typical situations for young drivers (Netherlands)
- Development of individual strategies for safe driving (Germany)
- Role plays (Germany)

Track module:

- The ' mickey mouse' exercise in the RACC training which focused on level 3 driver behaviour (effect of peer pressure, distractions, multi-tasking)
- The vehicle inspection in ECF France: pairs of participants examine each other's cars and draw conclusions about the state of the vehicle (maintenance and safety features) in relation to the accessories added (i.e. on what aspect of the vehicle is the money being spent- safety or aesthetics?).

[52] Large and hi-tech track training facilities may give the wrong impression of the training to the young drivers. They may increase expectations of a manoeuvring skills-based training. Such an impression may be strengthened by the presence of advertising or any other features related to racing, for example. The areas used in Spain included car parks which were adapted for safety reasons but which lead to no false expectations on the part of the participants.

8. ANALYSIS OF EVALUATION DESIGNS

This section looks at how the NovEV training programmes were evaluated, and how reliable and valid the results are. Different types of evaluation designs, and ways to implement them, were presented in the EU Advanced project report ('How to make a 5 star evaluation of your training'). Question scales were used by many of the participating countries from Hatakka M. (1998) Novice drivers' risk- and self-evaluations. Use of Questionnaires in Traffic Psychological Research[53].

8.1 Research designs

4 of the 6 evaluation designs (France, Netherlands and the two Spanish projects) were before-and-after measurements with control groups. The effects of the training were not measured immediately after the course; in fact they measured the effects up to 11 months after the first measurement.

In Germany, due to time constraints, only a process evaluation was carried out (single measurement with no control group). In Austria, a process evaluation was also carried out, in addition to a survey using before-and-after evaluation design (although the participants were not the same in each case: between subjects design) but without control group.

8.2 Objectives of the evaluations

The objective of 4 of the 6 evaluation designs (France, Netherlands and the two Spanish projects) was to measure changes in the knowledge, skills, attitudes and behaviour of the participants in the programmes. A full-scale evaluation of accident rates was not possible due to the short timeframe and small samples of novice drivers participating in the project. In Germany, the objective was to check the correct implementation of the programme which is being unveiled on a long-term nationwide basis. This was also Austria's objective. Another aim in Austria was to see how multiphase participants differed, if at all, from traditionally educated drivers 4 years previously.

8.3 Subjects

Aside from Austria, which has an obligatory multiphase system, participants in all other NovEV programmes enrolled on a voluntary basis. Subjects were selected on the basis of their age, sex and driving experience (or length of licence). Across all the participating countries, ages ranged from 18-24 and driving experience varied from 4 months to 3 years; in the case of one programme, minimum mileage was also required.

In Netherlands and France, participants were also chosen on the basis of their pre-licence education type. These countries wished to measure the effects of the 2nd phase training compared to the pre-licence education (RIS in the Netherlands; AAC in France).

The experimental and control groups were both chosen, using a random sampling technique, from the list of persons interested in participating in the training.

In Germany and the Netherlands, the views of the trainers (and trainers-of trainers in the case of Germany) were also gathered.

[53] Method Development, General Trends in Four Sample Materials, and Connections with Behaviour. 219p. Annales Universitatis Turkuensis, ser.B, Humaniora. Turku: Painosalama.

8.4 Data collection methods

The principal data collection methods were questionnaires and on-road driving audits.

Measurement instruments behaved partly in the predicted way, but not always.

Reliability measures showed that measurements were reliable, according to findings in this study or earlier in another context.

8.5 Methods used in analyses

The results and conclusions in the studies were based on statistical analysis and also partly on qualitative analysis.

8.6 Problems in reliability and validity

The drop-out rates were significant across the board, despite considerable potential rewards for those who participated in the whole intervention and in all the measurements. As a result, the statistical power in the analyses decreased. However, wherever possible, a repeated measures design was used in the analyses.

Except in Austria, the participants were volunteers. According to the programme, there were different motives for participating: road safety, reduction of the probationary period, winning a prize (e.g. a new car, free insurance for one year, a foreign holiday, etc).

Ultimately, the small group sizes and short follow-up period meant that there were no real possibilities to obtain results concerning actual safety effects in traffic (reduction in accidents).

On a detailed level, the evaluators of the feedback drives / driving audits often knew if the participant was in the experimental or control group.

Unfortunately, the difference in results and environments, and also evaluation designs, in terms of measurement methods as well as programme implementation, did not allow the results of each country to be combined. This was not an objective of the NovEV project, however.

Interestingly, the project showed that a difference between the original design and the training in practice, as well as differences in evaluations, seemed to produce differences in results. For example, small differences in the content and methods of track training may lead to significantly different effects on participants.

8.7 Conclusions

Despite the above comments concerning problems in the evaluations, the results can be considered reliable, mainly because of the before - after design with randomised experimental and control group.

9. OVERVIEW OF EVALUATION RESULTS

	Results	Evaluation design	Data source	Time of post-training measurement
Austria	Participants generally satisfied with obligatory modules Multiphase females reported safer driving and less speeding offences than traditionally educated females (4 years in between) Possible counterproductive message regarding track module	Before-and-after (no control group) Process evaluation	Questionnaires	na
Belgium	DROPPED OUT			
France	Significant positive change in awareness of risks linked to driving habits (MALES) Significant positive change in driving skills for defensive driving (MALES) (Slight) trend towards less frequent risky driving situations (MALES)	Before-and-after with control group	Questionnaire Questionnaire Questionnaire	7 months after 2nd training day
Germany	Programme mostly implemented as intended Training rated positively by participants Participants claim to use practical driving strategies developed in training Possible counterproductive understanding of the goals of track module	Process evaluation (single measurement)	Questionnaires	na
Netherlands	Significant positive change in calibration in 1 of 2 test groups Significant positive change in driving skills* in 1 of 2 test groups Negative trend in driving skills* and calibration in other test group *Driving skills = GDE levels 1 and 2	Before-and-after with control group	On-road audit On-road audit On-road audit	1 month after training
Spain RACC	Significant positive change in social driving behaviour	Before-and-after with control group	Questionnaire	6 months after training
Spain RACE	Significant positive change in knowledge Significant positive change in driving skills (GDE levels 1 and 2)	Before-and-after with control group	Questionnaire On-road audit	6 months after training

10. CONCLUSIONS

The following chapter focuses on some general conclusions, followed by conclusions on specific levels of the programme: the design level, organisational level and content level.

10.1 General conclusions

Firstly, it should be remembered that the sole focus of the NovEV project was one or more days of road safety training after the licence. This training was simply added to whatever basic training the participants received. It is vital, however, to stress the importance of a well-developed basic training too (see EU BASIC project report, 2003). Furthermore, the Finnish experience suggests that the pre- and post-licence training should be linked in some way (common messages, etc).

The results of the 2nd phase pilot programmes in the NovEV project (see previous page) show that such programmes can have a positive influence on the –mostly self-reported - driving behaviour of novice drivers, as well as, in one case, trainer-audited driving behaviour. This, at least to some extent, confirms the validity of the guidelines laid down in the Advanced project on 2nd phase training. Importantly, however, the results also show that it is quite possible for 2nd phase programmes to have a negative impact on the driving behaviour of such drivers. They also show that, despite the best efforts of the designers of the programme, novice drivers can receive unintended signals about what the course is actually supposed to achieve. These points, again, confirm the concerns expressed in the Advanced project.

All programmes with a comprehensive before-and-after evaluation design with control group succeeded in achieving positive results in the NovEV project. Although the results were perhaps not as positive as hoped for –a number of other factors were measured and no significant results were achieved -, the NovEV project should nevertheless be considered a success. The participating countries succeeded, at least on paper, in creating 2nd phase training programmes based on the guidelines established in the Advanced project. These guidelines stressed the importance of focusing on the higher levels of driver behaviour, and using participant-centred methods designed to generate discussion and self-reflection. However, the results were by no means an overwhelming success, indicating that further work needs to be undertaken to verify the assumptions of the Advanced guidelines. It should be added, moreover, that there is no clear evidence at this stage proving the effectiveness of existing obligatory 2-phase systems, such as in Finland, Luxembourg, and more recently, Austria (see annex 5 on the Finnish system).

Translating the Advanced guidelines from paper into practice poses a number of potential problems. Firstly, the managers of the training organisations must agree on the philosophy put forward in the guidelines. This project has highlighted the fact that organisations may still be reluctant to do so[54]. Clearly, more effort needs to be spent communicating the philosophy of Advanced to trainers and training organisations who traditionally have had a different approach. Secondly, the trainers who are responsible for implementing the course must be properly trained. Training-of-trainers needs not only to focus on the goals and implementation of the training, but also on how to deliver the course in the right way. The trainers need coaching skills, knowledge of young driver psychology and experience with group dynamics. These skills

[54] In particular, track trainers with experience in voluntary post-licence training may have difficulty accepting and implementing 2nd phase guidelines (as stated in the Advanced report). This difficulty may relate to their acceptance of the rationale of the 2nd phase guidelines as much as to a tendency to fall back on their previous working methods. It also depends on the quality of their training.

do not come overnight and need to be developed and honed over time. Thirdly, despite the best efforts of designers, training organisations and trainers, it is still possible for the novice driver participants to draw the wrong conclusions about the training. This phenomenon is particularly relevant to the track modules where participants may be left with a feeling that they have improved their mastery of emergency traffic situations, even though such mastery is not an objective of the course. As we know, such feelings can easily generate over-confidence amongst some young drivers, with potentially disastrous results when driving independently. The above points all lead to the same conclusion – a 2nd phase programme must be scientifically evaluated at all these levels to ensure that it is being implemented in the way it was intended. As seen in the Dutch NovEV experience, the same training on paper had completely different results in the two training locations.

The difficulty of translating the guidelines from paper into practice may, to some extent, explain why the overall project results were less comprehensive than originally hoped for. Other more structural factors may also explain this. It is commonly agreed that a 2nd phase programme should be spread out over time. This lengthens the support period for novice drivers and allows them to acquire more experiences and to implement what has been learned in practice. In the case of the NovEV projects, however, 3 of the 4 programmes (where changes were measured scientifically) took place over only 1 day. A single day's training is unlikely to lead to significant changes in the attitudes, skills, knowledge or driving behaviour of the participants. In addition, the post-training measurement phase in 3 of the 4 programmes took place 6-7 months after the training. Therefore, it is possible that the impact of the training had already begun to fade by that stage. That said, one of the objectives of this type of training is to encourage the development of sufficient self-evaluation skills for the effects to be longer-lasting. Another possibly explanatory factor is that the novice drivers participated on a voluntary basis. It is likely that such participants were already more safety-aware than the average novice driver before the training began. Logically, therefore, it would be more difficult to raise their (already high) level of safety awareness in such a short period of time.

Another important conclusion of the NovEV project is that the results – in terms of training effects on participants - can only really apply to the group-types represented in the various training programmes. Despite material incentives designed to attract a representative spectrum of the young and novice driver population, it is likely that the final sample groups were, in general, more safety-oriented than in the overall young driver population.

What is also clear from the experiences of the pilot programmes within NovEV is that novice drivers are not interested in participating in 2nd phase programmes on a voluntary basis. Despite a wide range of incentives to take part, all pilot countries had 1) difficulty in convincing novice drivers to participate, and 2) a major drop-out rate once the training had begun. Positively, however, the NovEV programmes were all - with the exception of the one training organisation where the negative effects were recorded - highly rated by the final participants. Moreover, the Austrian findings support the data collected over the last few years in Finland: namely that, although novice drivers may not be that keen on the idea itself, they are highly satisfied with the obligatory 2nd phase training once they are there.

10.2 Conclusions at different process levels

10.2.1 Conclusions: design level

- Trainers can vary, in terms of their background and profiles, from country to country. In Finland, for example, the driving instructor is responsible for all 2nd phase modules. In contrast, a different trainer is present for each training module in Austria (track trainer, instructor for the feedback drive and a psychologist for the group discussion). It should be noted that trainers with only a short period of contact time with participants have a tendency to try to transfer their knowledge and expertise to the young drivers. This should be avoided, because of the role of the trainer in second phase as a coach, not an instructor. Trainers are responsible for raising the right questions, not the right answers. On the other hand, should driving instructors, for instance, be given the task of ensuring the implementation of the whole 2nd phase training? Coaching skills, and to some extent experience with groups, are not skills typically associated with driving instructors. In summary, this is an ongoing discussion point, but ultimately, whatever the choice is made, the trainer should be able to deliver the course properly.
- The vocabulary used in training is very important, because the words themselves send messages to the participants. Word such as 'instructors', 'track training' and 'trainers' can all reinforce the impression that the objective of the 2nd phase is to improve manoeuvring skills, through a process of repetition and measurable 'improvement', until a fixed objective is reached (and that the programme is based on following instructions rather than thinking for oneself and reaching individual conclusions). Alternative, more neutral (and more accurate) wording could be sought after. For instance, track training might become 'driving demonstrations', and trainers and instructors could become coaches.

10.2.2 Conclusions: organisational level

- Trainers require training over a sustained period of time, rehearsals, feedback and ongoing training. Accreditation of trainers should be seriously considered when implementing 2nd phase training on a national level. Apart from the obvious need for trainers to understand and deliver the intended content of the training, trainers require coaching skills, an understanding of young driver psychology and the ability to deal with - and generate discussion in - groups.
- Training should, where possible, allow the participants and trainers to relax in each other's presence before the group discussion module takes place. (Youngsters may be less inclined to talk openly at the beginning of a training day). In Finland, for instance, the group discussion now takes place after the track training. This provides for the above conditions, at the same time as allowing for the experiences from the track training to be combined with their independent driving experience in the discussion.
- Participants should be monitored, perhaps through questionnaires, to see what conclusions they are drawing from the training. This is especially relevant to track modules where participants may draw the wrong conclusions, thereby signalling to the training management that the content or delivery of the training should be changed.
- Some form of shelter should be available to groups during track training sessions (in the event of extreme weather conditions). As stressed in the Advanced report, the discussion sessions following track exercises are extremely important, and the effectiveness of such discussions can be impaired if the participants are unable to concentrate properly.

10.2.3 Conclusions: content level

- A feedback drive should be a feedback drive, not a driving lesson. Furthermore, the feedback should be mostly coming from the participant, not the trainer. The trainer should evoke feedback from them by asking targeted questions. It is important to emphasise and monitor this in the training-of-trainers and the rehearsal. It is very easy for trainers to tell participants what the conclusions are of the exercise, rather than to ask about the participants' experiences and to get participants to conclude for themselves what they will take from the exercise and the course. The presence of other participants in the car (taking it in turns to drive and commenting on each other's driving styles) can be very beneficial. This is not just from the perspective of gaining peer feedback but also because it allows for discussion on how the presence of passengers can influence one's driving style (e.g. peer pressure). Ample time should be allowed for each participant to relax and to drive 'normally'. Enough time should also be allocated for a discussion at the end of the session.
- Track modules should focus primarily on risk awareness, not manoeuvring skills training[55]. Unnecessary repetition should be avoided as this reinforces the impression that skills are being trained. Discussion should take place following each exercise. Careful thought should be given to the location of the participants during the exercises – to ensure a full learning experience (should they drive themselves, be a passenger, be standing outside the car, have a demonstration from the trainer?). Risk awareness exercises training the higher levels of driver behaviour are possible in track modules and are to be encouraged. The amount of time spent in the track session(s) should not be disproportionate to the time spent in the feedback drive(s) and group discussion(s). The results of the subjects' satisfaction questionnaires in Austria and Germany, and the results of the Netherlands' track training support these findings.
- Group discussion should be primarily a discussion, not a lecture. The trainer is required to pose questions rather than inform, to guide the discussion rather than lead it. Participants should be analysing experiences and engaging their brains. Participants should be encouraged to identify areas of risk and to relate their own driving to the situations evoked. Feedback from the participants should be written on a flipchart. This is a neutral form of registering comments and is less confronting than direct conversations between individuals. Discussions amongst the group should remain low-key and relaxed. Again, the main focus should be on the higher levels of driving behaviour. Videos, case studies and role plays can provide considerable structure to these discussions. Such structure also helps trainers who are not that confident or experienced. Videos or screen presentations should not be overly relied upon, however, as they can easily become another form of presentation / lecturing. The starting point, where possible, should be the experiences of the participants.
- The period between training days can also be structured to good effect, in order to encourage young drivers to be aware of and to analyse their experiences. Self-evaluation forms and driver profiling can aid this process and can mentally prepare the drivers for a forthcoming training module. (This is already done in practice in Finland).

[55] Emergency braking practice is an exception to this rule due to the importance of the skill and the relative simple manoeuvre itself, compared to braking and avoidance for itself. Emergency braking practice should either be designed to improve emergency braking skills OR to encourage risk awareness (stopping distances in relation to speed and surface, reaction times, etc). If the objectives of the exercise are mixed, i.e. to improve both the braking manoeuvre itself (=skills training) and to improve risk awareness, the young drivers tend to remember the skills element and the risk awareness message is lost.

11. RECOMMENDATIONS FOR 2ND PHASE TRAINING

Based on the conclusions in the previous chapter, and on prior experience in the field of 2nd phase and 2-phase driver training systems, the following recommendations apply. These recommendations should also be considered in conjunction with those already stated in the EU Advanced project. This chapter is written in the same manner as the conclusions, with recommendations categorised under 'design', 'management' and 'content'.

11.1 Design of 2nd phase

- ✓ The content of the 2nd phase should focus primarily on the higher levels of driver behaviour and should be delivered using participant-centred methods designed to generate discussion, self-evaluation and the drawing of individual conclusions and strategies for safe driving.
- ✓ The programme should be spread out over time, in an effort to support the novice driver in a structured manner through his/her early independent driving experiences, and to maximise the potential for behavioural change.
- ✓ NovEV recommends the implementation of 2nd phase training during the first year of independent driving following the licence. This timing takes into account both the very high risk levels of novice drivers immediately after the driving test AND the need for these drivers to have some sort of practical independent driving experience before attending the 2nd phase. The modules of the 2nd phase should be spread out during this period (as in Austria), in order to offer an ongoing support mechanism for the novice driver.
- ✓ The content / messages of the 2nd phase training should be linked where possible to the pre-licence training (in order to ensure coherence in the training and to reinforce the training objectives).
- ✓ The training modules (class, track and road sessions) should be balanced in terms of length and focus. A disproportionate emphasis of one training module, particularly track training, risks sending the wrong message to the novice drivers.
- ✓ Track modules should focus primarily – if not exclusively - on risk awareness and should strive to avoid unnecessary and counterproductive emphasis on vehicle control skills. Track exercises designed to simulate situations involving the higher levels of driver behaviour are to be encouraged.
- ✓ For political, scientific and logistical reasons, countries may like to consider the possibility of phasing-in the different modules of an obligatory 2nd phase programme over a period of time. The rationale of this procedure is to start initially with a feedback drive (for which the driving instructors would need to be trained). Novice drivers during this period would only have the feedback drive as their 2nd phase programme. At some stage later in the future, once enough training has been given to the trainers, a group discussion module could be implemented too. Then, finally, a track module could be introduced at a later date. This phasing-in process would achieve several goals. Scientifically, it would be possible to measure and isolate the influence of one or a combination of different modules from each other. Politically, it would be less drastic than introducing a 3-module programme at the same time. Logistically, it would also provide time for the proper training-of-trainers, rehearsals and controls before the (most complex) individual modules (group discussion, track experiences) are introduced.

11.2 Management of 2nd phase

- ✓ Detailed guidelines for managers and trainers should be made available in the event of 2nd phase programmes being implemented. Any law allowing for the implementation of the 2nd phase will be general and will not provide the level of detail necessary to ensure a coherent and uniform training across the country (and via different training organisations and trainers). A detailed 2nd phase training manual should be written and distributed thereafter, outlining the objectives, content, methods and process of the programme. Individual trainers' manuals should also be developed for each specific training modules.
- ✓ The implementation of the 2nd phase programme should be quality-monitored and managed by an independent steering committee. Whereas organisations and individuals involved in the 2nd phase training may have a consultative role, decision-making should clearly be in the hands of independent persons. This steering committee should be responsible for ensuring adequate training of trainers, accreditation of individual training bodies and trainers, and for ongoing monitoring of training over time. Whilst the exact content of the specific training modules may vary from one place to another, the content must be designed to reach the objectives of the 2nd phase programme, and any exercises deviating from norms laid down in the official training manuals must be approved by the independent body. In the knowledge that there can easily be a difference between the design on paper and implementation in practice, ongoing, independent quality control, conducted by trained auditors, is vital.
- ✓ Trainers require training and testing on coaching and moderating groups of novice drivers. Despite the existence of alternative options (psychologists, sociologists, social workers…), the obvious choice of 2nd phase trainer (particularly for the class and feedback drive modules) is the normal pre-licence driving instructor. Whereas in some countries, for example Germany, driving instructors are trained and experienced in moderating young groups of drivers, instructors in many other countries receive neither training nor do they have the opportunity to practise this skill. As coaching and group dynamics are so important for 2nd phase training, countries where the trainers have little or no prior experience in these areas should think twice before implementing obligatory post-licence training. NovEV recommends that the European Commission should consider the benefits of a new EU project to design and test coaching and moderation training seminars for driving instructors.

11.3 Content of Training Programme / Trainers

- ✓ More examples of specific training exercises should be made available to countries and organisations who wish to implement 2nd phase training. Current 2nd phase guidelines remain largely theoretical at this stage, so more practical examples would be beneficial, and would go further to ensure that the training is carried out effectively and coherently. NovEV recommends a follow-up European project to collect and create effective and innovative exercises for 2nd phase training (track, class and on-road), building on the examples already provided in the Advanced project risk awareness database. These examples should focus above all on levels 3 and 4 of the GDE matrix.
- ✓ Monitoring is required to check the perceptions of both the track trainers and the novice drivers with regard to the messages/goals of the track training. At least two NovEV countries reported participants feeling more skilled (in terms of vehicle control) as a result of the training, despite this not being an objective of the track training. (This observation is linked to the potentially counter-productive phenomenon of overconfidence amongst novice drivers).

ANNEXES