Evaluation of C-ITS and Automated Driving

A cyclic approach in five parts

Henk Taale, Hans van Lint, Isabel Wilmink
DUTCH ROUND TABLES FOR SMART MOBILITY | RONDETAFELS.DITCM.EU
ROUND TABLE EFFECTS
22 NOVEMBER 2016

1 Introduction

The cost-effectiveness of traffic management is an important topic, which draws more and more attention on all levels: from policy to operations. There is a growing need to know the effects of measures or programmes in advance or after implementation. Not only to be transparent and to account for investments done, but also to learn from real-life experience and to come to a better implementation or tuning of measures in the operational environment and for the benefit of a more effective policy and useful future investments. In this way the plan-do-check-act cycle is closed.

Evaluation can be classified on several aspects. A first distinction is between ex ante and ex post evaluation. An ex ante evaluation is focussed on the feasibility of measures and several methods can be used for that: literature review, expert judgement or a modelling study. Especially a modelling study has advantages: it is cheaper than a real-life pilot, it has no impact on drivers or safety and with a model it is relatively easy to study alternatives. An ex post evaluation is carried out after the implementation of a measure and typically consists of a before and after situation for which data is collected and analysed.

Another distinction can be made between certain types of evaluation:

- A technical analysis: does the system perform according to specifications?
- An impact analysis: what are the effects on traffic, safety or the environment?
- A socio-economic evaluation: what are the costs and what is the effectiveness related to these costs?
- An analysis of legal and institutional aspects: are privacy issues involved and how can they be dealt with?
- An analysis of behaviour and public acceptance: do the road users react as expected and are the willing to follow the instructions given?

There are more classifications possible, but we restrict ourselves to the two mentioned above. These different distinctions are not only relevant for traditional traffic management, but certainly also for the evaluation of C-ITS and automated driving. The question is how we can use these evaluation types to determine the effects of C-ITS and automated driving. To answer this question we first look at existing evaluation methods and then turn our attention to an attempt to integrate evaluation as a common thread in the development of C-ITS and automated driving.

2 Evaluation methods and aspects

Since long there is an interest in learning from experiments and their results. To be able to do that good evaluation approaches are essential. For the evaluation of C-ITS and automated driving we tried to find a satisfying approach. An inventory of existing evaluation methods was conducted and the methods were assessed on their suitability to evaluate C-ITS [1].

2.1 Evaluation methods

The assessment was done for a number of Dutch approaches as well as European, among which the FESTA Handbook is the most important one [2]. There are other ones, but most literature related to the evaluation of C-ITS refers to or uses the FESTA approach. For the assessment a number of aspects was investigated, both aspects specific for C-ITS as well as general aspects. Specific C-ITS aspects had to do how the method deals with the level of impact (vehicle, traffic operations, network or societal), data collection and if the design of the method is able to take into account C-ITS. More general aspects had to do with data collection, KPI's, tools, documentation and completeness of the method and the expertise and effort needed to apply it.

From the inventory it was concluded that all methods have their strengths and weaknesses and that none of the method contains a complete approach. All methods emphasis a timely involvement of evaluation expertise in a project, but they do not or only limited address their applicability. Most suited for the evaluation of C-ITS and automated driving is the FESTA method. Its steps give a reasonable complete picture of what is needed. However, it does not cover the whole chain from policy to operation and back. Therefore, deepening is needed for three reasons: the problem of scaling results to a higher level (from vehicle to traffic flow to network), the translation from and to policy goals and closing the 'evaluation circle', that is using results to ask new questions (acceleration of R&D) and/or to formulate new policy goals.

In the following this is described in more detail. First, the aspects for evaluating C-ITS and automated driving are addressed and after that the evaluation circle is described. This includes the issues which are mentioned above.

2.2 Aspects for evaluating C-ITS

There are several aspects that require attention when evaluating C-ITS:

- In many cases **in-vehicle measurements** are needed, for instance to evaluate usage, compliance with advice and driver behaviour. Depending on the application (and size of the trial/implementation), it can be important to be able to distinguish whether effects are the consequence of the driver using the system or because of other circumstances (e.g.: does the driver slow down because of a reduced speed advice, or because of a slower predecessor?).
- Because in-vehicle measurements are needed, the **penetration rate** is an important factor. In many cases, penetration rates will initially be low, or will be low for a long time to come. Many interesting indicators can then still be determined using data from vehicles, but impacts on the traffic flow level cannot be observed. If any impact depends on communication with other equipped vehicles, high penetration rates are required for any evaluation. The desired penetration rate probably needs to be orchestrated in that case. This means that the penetration rate is an important factor that determines the ratio between the measured effects and the estimated impacts. An higher penetration rate means that more effects can be measured and less have to be estimated.
- Only if penetration rates are (very) high, and a clear impact on driver/vehicle behaviour can be expected, is it possible to measure impacts with road-side sensors or 'generic' probe vehicle data. Even with high numbers of equipped vehicles (high fleet penetration rate), the **concentration in time and space of usage** is not necessarily enough to have high penetration rates at locations and during periods that are interesting for the evaluation. With these variable penetration rates also the relation between the effect and influencing factors and conditions varies.
- Impacts may occur only during very short time periods, in very dynamic conditions. It is interesting to measure what, at that time, the **interaction was with other vehicles** (the direct environment of the vehicle). This may require special measurement equipment, for example cameras, but it could give insight into a number of issues. In this context an interesting question is whether the interactions between equipped and non-equipped (non-communicating) vehicles is different from the interactions between equipped vehicles, or the interactions between non-equipped vehicles. In that case: Can the behaviour of non-equipped vehicles be witnessed / reconstructed? Was their behaviour a response to the behaviour of equipped vehicles (perhaps

behaving notably differently than non-equipped vehicles?), or are there other causes for their behaviour?

- Several in-vehicle applications are often combined in one 'package'. What can then be said about **separate effects or combined effects**? For the evaluation it is important to log which application is active when or if several applications were active at the same time. And if so, what this meant for the advice given or for direct interventions in the vehicle's behaviour. If effects are measured separately, it should be noted that the joint effect of multiple applications is not necessarily the summation of all separate effects – one application may take away some of the effect of another application, or one application may add to the effect of another one. For instance, a forward collision warning system may reduce the number of times an emergency braking systems needs to be activated.
- **Scaling up of effects** found for an application under certain trial conditions to a higher level (higher penetration rates, all traffic conditions, larger road network) often requires additional information on a range of situational variables, many of which are difficult to measure with the vehicle or difficult to obtain from another source. If possible, situational variables should be logged or at least estimated. Simulations in a traffic model may be needed to predict the effect in other networks or under different traffic conditions. If possible, the modelling should include the level of compliance found in real-world situations (and any variations in compliance depending on the situation), and changes in the interactions between vehicles.
- If simulation tools are used for scaling up (of for exploring effects in conditions not covered by a field trial), it is important to verify that the tool used can, in fact, produce the relevant behaviour of vehicles. For the current situation this can be validated, but for the situation with C-ITS there is not much knowledge about the behaviour. Therefore, the **suitability of the simulations tools** may be limited, or may vary considerably depending on the type of application or the traffic conditions considered.

2.3 Aspects for evaluating Automated Driving

Many of the aspects mentioned for evaluating C-ITS applications also hold for automated driving. However, there are a few additional aspects that need to be taken into account:

- **Automated vehicles are not easy to obtain** for testing (especially ones of SAE-level 2 and higher).
- Not many researchers (evaluators) have ever driven an automated vehicle (or been a passenger in one). Not much information (if any at all) is available from manufacturers or independent testers. This means that very **little is known about the behaviour of automated vehicles in real-world traffic**. This concerns most of all their lateral behaviour (lane changing – especially at motorway entrances and weaving sections). But basically, there is very little information about any choice a human driver would make (from route choice to, say, speed choice under varying weather conditions). This makes it difficult to decide on what should be measured under what kind of conditions, if a field trial were to be set up.
- There is much speculation about whether automated driving will lead to **more trips** (because drivers can do something else and will travel further, or because of empty vehicles).
- Even more so than with C-ITS applications, it is important to learn about the interaction between 'equipped' and 'non-equipped' vehicles (**interaction between automated and manually driven vehicles**). It is clear that automation may sometimes lead to 'unnatural' driving behaviour which may lead to particular reactions from other vehicles (mostly the manually driven ones, but possible also other automated vehicles).

- Again, **simulation is an important tool for exploring the effects of automated driving** (benefits and disbenefits). However, knowing little about the behaviour of automated vehicles means that many assumptions have to be made. **No ground truths are available**, so models cannot be validated and even verifying models can be difficult. It has to be considered that automated vehicles of different manufacturers (and levels) will behave differently, so just modelling one type of automated vehicle does not suffice. Also, variation in behaviour of manually driven vehicles must be incorporated. Some spectacularly wrong assumptions about automated vehicle behaviour have already been observed in some studies (leading to unrealistic expectations of benefits), so care must be taken to document exactly what has been modelled. Lateral behaviour especially should receive a lot of attention.

3 Evaluation Circle

In the previous paragraphs we have described existing evaluation methods and important aspects for the evaluation of C-ITS and automated driving. We now turn our attention to the integration of all these issues. For this purpose the evaluation circle was developed [3][4]. The circle uses the principles from the evaluation methods described and adds an iterative approach using ex ante and ex post approaches as shown in figure 1.

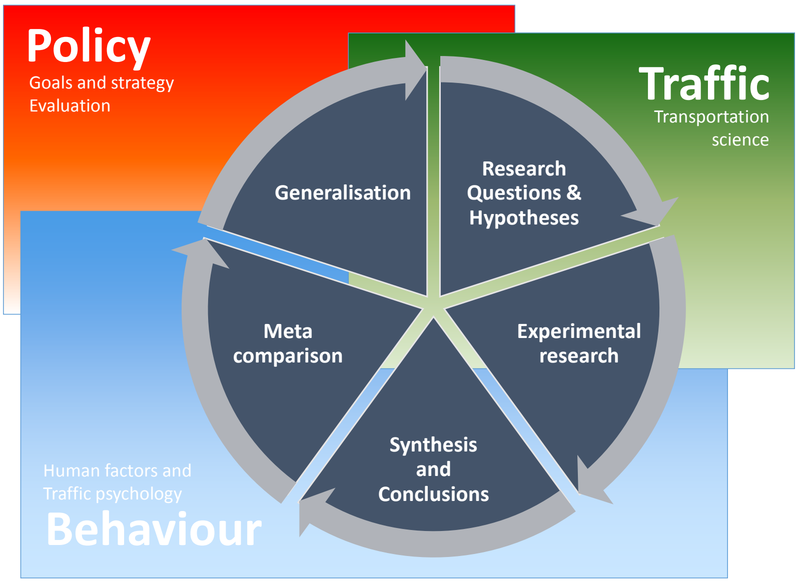


Figure 1: Evaluation circle

The principle behind the evaluation circle is very simple: evaluation is a continuous process, which does not stop when an evaluation report of a pilot study is written. Generally speaking, the first three steps (from hypotheses to experimental research to conclusions) involve ex post evaluation of C-ITS and automated driving using either field experimentation and/or empirical research using for example driving simulators or surveys. The objective of these steps is to unravel the functioning of C-ITS and automated driving in terms of driving behaviour, technology, user acceptance, etc. Typically, these steps are part of the many pilot studies done today. However, the final two steps (meta comparison and generalisation) are typically NOT incorporated in pilots today and are in our view fundamental to pave the way for large-scale implementation and technology take-up. Since the efficiency, safety and other direct or indirect effects of C-ITS are a result of interactions between drivers in vehicles with and without new technologies and not all general information about these interaction is available from local experiments, we need to interpret results in the light of the state-of-the-art worldwide (meta comparison); and use simulation as a means to predict ex ante what will happen under larger penetration rates (generalisation). In doing so, we can look at worst- and best-case scenarios and identify which critical design issues (related to technology and methodology) we need to address in the next round of pilot studies. Most importantly:

A cyclic approach to evaluation enables systematic programming of R&D in large programs (like in The Netherlands Optimising Use and SmartwayZ.NL), tailored ultimately to large-scale deployment of C-ITS and automated driving. Focusing on just pilots and technological realization leads to myopic R&D efforts in which there is a real risk of either ignoring or even completely missing valuable insights and lessons learnt. Systematic evaluation enables failures to function as drivers for progress and collective learning, and success as a means to build technical and methodological capacity and to raise support for standardization.

The challenge is that this cyclic approach by nature requires is a multi-disciplinary approach. To evaluate C-ITS and automated driving several domains are involved. In figure 1 three domains are shown: the policy domain (which includes decision making), the domain of transportation science (traffic and transport) and the domain of human factors and traffic psychology (driving and travel behaviour). Figure 1 shows that these domains are strongly overlapping. There are of course many more disciplines involved such as computer and data science; mathematics; control and automation; communication; mechanical and automotive engineering to name just a few. In the following we focus on these three core domains and the five parts which form the circle.

3.1 Domains and their parts

In formulating policy goals and in implementing traffic and transport strategies questions arise about feasibility; effectivity; societal benefits and costs; and how these must be judged and compared. These questions can be addressed in experiments. If these experiments lead to viable results and conclusions, combining them with results from other studies leads to valuable information, which can be used to evaluate or reformulate policy goals and strategies. Clearly, this is possible only if the results and conclusions can be generalised into recommendations for policies beyond a local trial or pilot.

Traffic and transportation science is the domain which studies the dynamics of transport and traffic systems over both short and long time scales (from operations to planning) on roads, corridors and entire networks. This knowledge is operationalized in mathematics, methods, and (simulation) models, and is fundamental to design the experiments and on the other hand to interpret the outcomes. Traffic flow theory studies interactions between vehicles and traffic simulation models use coarse behavioural assumptions to describe these interactions either macroscopically or microscopically. To evaluate C-ITS and automated driving these coarse assumptions are not sufficient, because the interactions take place on a more detailed level.

Drivers with technology will change their behaviour, but so will drivers without technology. The behavioural sciences play an important role in understanding and predicting how these behaviours will change and how we need to adjust our models to make sensible predictions of these future interactions. Moreover, behavioural sciences are important also in the design of human-machine interfaces, and in understanding compliance and user acceptance of technology.

3.2 Research questions and hypotheses

Every research starts with one or more questions. Normally these questions stem from the need to know something in relation to the policy at hand. It is also possible that questions arise from previous research. In all cases, they should relate to the measure(s) or system(s) under investigation. The nature of the questions can vary between very detailed, such as 'does this in-car system dampens shock waves?' or 'how does the driver react to this specific warning signal?', to very global questions, such as 'what is the impact of automated driving on the capacity of the road?' or 'which penetration rate of this C-ITS system is needed to have a positive impact on the traffic flow?'. In short: What do you want to know or to learn?

To be able to formulate good research questions it is important to take into account the circumstances in which the measure or system is supposed to have an effect. In evaluation terms we call this the **use case**[1]. On the other hand there are also circumstances which may affect the effectiveness of the measure or system which are beyond their control (e.g. weather conditions). That is called the **situation**. When a system or measure is evaluated, it is always in the context of a **scenario**, which is a use case in a specific situation and hypotheses should reflect that. A **hypothesis** is a statement about the effects of a system or measure on a certain traffic or behaviour characteristic (indicator), which is valid for a certain scenario. A hypothesis should always be phrased in terms that can be made S.M.A.R.T.: specific, measurable, acceptable, realistic and timely. In figure 2 the relation between these concepts is shown.

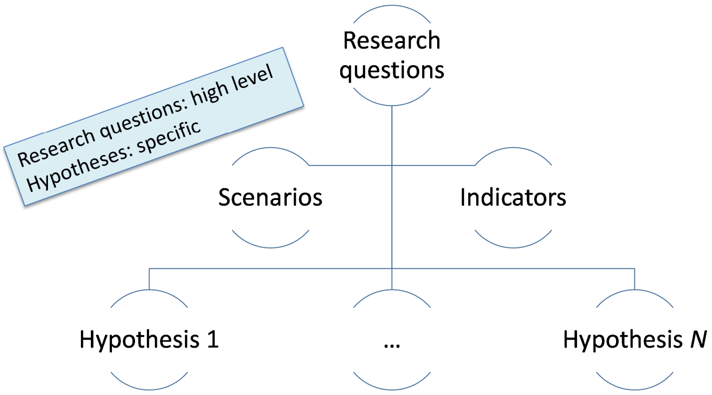


Figure 2: From research questions to hypotheses

Definitions

Use case:	circumstances in which the measure is supposed to have an effect and in which the user should show a certain behaviour
Situation:	circumstances which may affect the effectiveness (beyond the control of the measure)
Scenario:	a use case in a specific situation
Hypothesis:	hypothetic answer to research question under scenario in terms of indicator(s)

3.3 Experimental research

Experimental research is not only about impact. For example, also technical tests or cost-benefit analyses could be part of an experiment. In any case, experiments have to be designed carefully. A good design is crucial and determines if an answer to the research questions can be found in the first place. Among others the research design should cover the aspects from data collection to reporting results. The FESTA approach is very useful for this. The steps it covers are shown in figure 3 and are described in detail in [2].

[1] Note that this definition is different from what normally is referred as a use case in the C-ITS domain.

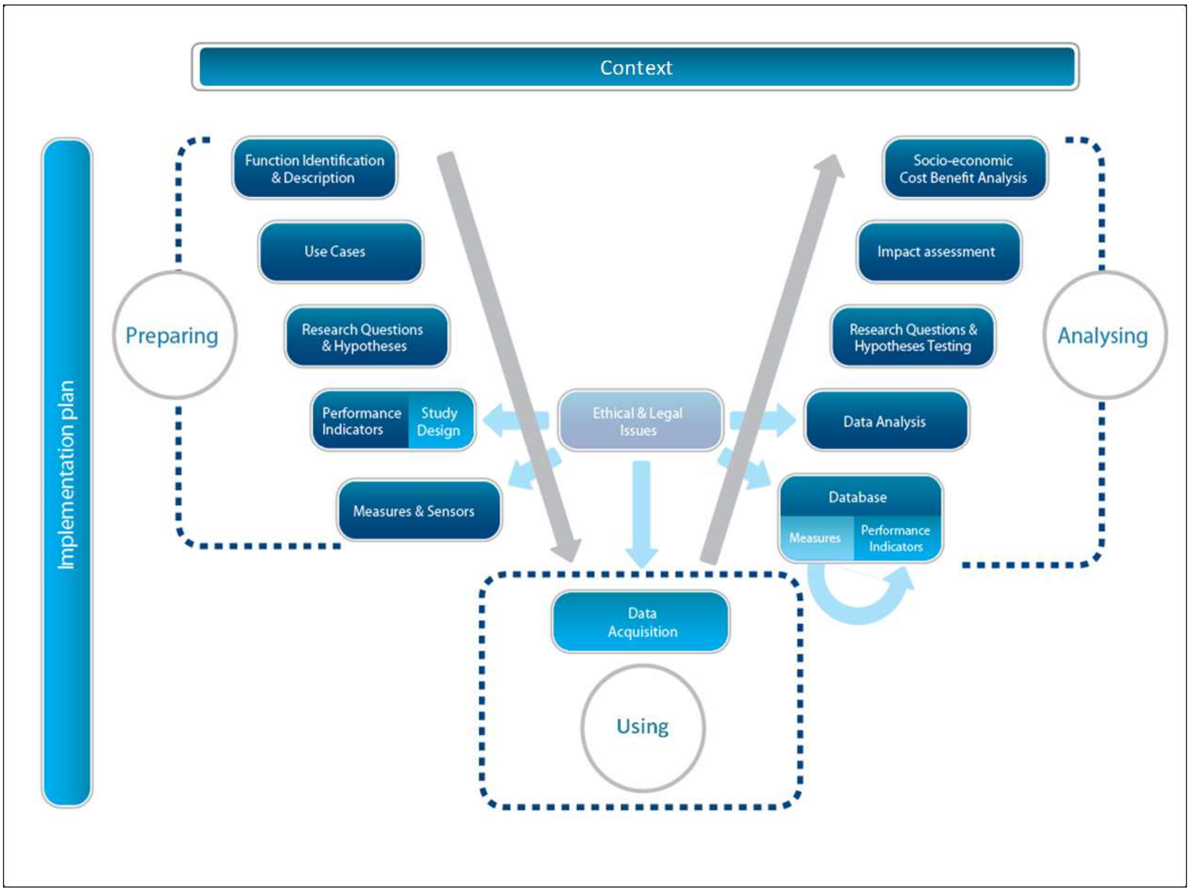


Figure 3: Steps in the FESTA Handbook [2]

As shown in the figure the steps in the FESTA handbook also cover aspects as described earlier, such as defining use cases and formulating research questions and hypotheses. A lot of attention is paid to data: measuring, collecting, storing and analysing data are described in detail, including a lot of practical aspects that are important to take into account when an experiment is conducted.

As for the evaluation design and the accompanying data acquisition methods a lot of different approaches are possible. Traditional traffic management was mostly evaluated using the simple before-after scheme. Data was collected during these two periods and was compared using the proposed indicators. Sometimes an alternating scheme (off-on-off-on …) was used, but evaluations were relatively straightforward. For C-ITS and automated driving things are a little bit more complicated, as was explained in paragraph 2. To understand the challenges, consider figure 4, which places different data collection techniques along two dimensions.

The vertical axis indicates whether the data actually reflects human behaviour (travel and driving choices, interaction, compliance, etc.), whereas the horizontal axis indicates the degree of experimental control under which the data can be collected.

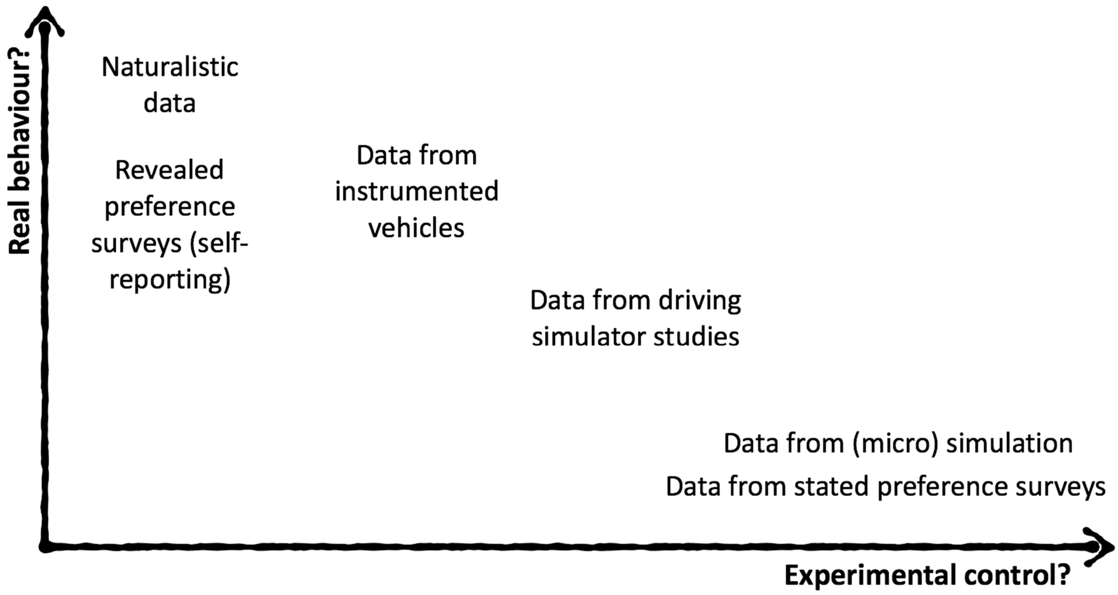


Figure 4: Data collection - trade-off between validity and experimental control

- It is possible to simply ask road users what their behaviour would be in certain situations (**stated preference surveying**). With these explorative inquiries experimental control is optimal and a rich and thorough body of knowledge is available to design SP surveys and estimate statistics and models from these. However, there are large limitations in terms of the validity of the outcomes that relate to for example representativeness, the lack of actual context (e.g. physical discomfort or sense of urgency) and bias towards social acceptable answers.
- **Simulation** (on a computer) can be used to get ex ante insight in the possible effects. Microscopic simulation can be used to determine the impact of different driving behaviours under C-ITS and VA and provide an ideal safe artificial test bed. Clearly, simulation gives full experimental control over the scenario's and use cases considered – experiments can be repeated hundreds of times while varying only one or two variables and full sensitivity analyses can be done. However, most of the commonly used models are not valid yet to simulate C-ITS or automated driving [5] for a number of reasons. Firstly, fully automated vehicles in a sense are easy to simulate, but C-ITS, and partial automation may change driving behaviour far beyond the validity of existing car-following and lane-change models in microscopic models, particularly in terms of the response to information and control, the process of authority transition and many other aspects. Secondly, the response of non-equipped vehicles to automated vehicles is an as of yet poorly researched area. We have very little idea how "other vehicles" will interact with partial automated vehicles. Thirdly, for C-ITS and automated driving also the communication between vehicles and the roadside and between vehicles themselves should be emulated, particularly for worst-case scenarios (what if communication fails?). Notwithstanding these validity problems, microscopic simulation models do allow us to explore the possible range of effects and impacts (worst, best case) and ask what-if questions (see further below under "Generalisation").
- Experiments with a **driving simulator** let *real* people drive in a controlled (simulated) environment. For obvious (ethical) reasons, driving simulators are the only viable experimental platform to use researching distraction, risk taking or safety related hypotheses in general. But also in many other cases, driving simulator studies provide a rich experimental environment. Driving behaviour can be observed in detail, while monitoring also underlying mental and physiological processes. By combining driving simulator studies with surveys and interviews (e.g.

about their decisions and their use of information) a wealth of detailed data becomes available. However, typical validity issues are the often small samples, the representativeness of the drivers in the experiments (bias towards students), the level of realism of the environment presented to the drivers and the validity of the simulator environment to research particular types of behaviour (car following, gap acceptance, speed choice, etc.).

- To increase the level of realism, an **instrumented vehicle** can be used, in which drivers can be instructed to drive according certain strategies or simply as they would do normally. Due to the observation systems in the vehicle the same level of detail can be obtained as with a driving simulator, and drivers can of course also be interviewed. Although we now have a real driver in a real vehicle under real-world conditions the presence of observation systems or an observer in the vehicle could lead to biased results and limit validity. Instrumented vehicles offer less experimental control than driving simulator studies due to a more limited range of (repeatable) circumstances and traffic conditions that can be used. The aforementioned ethical (safety) issues furthermore constraints the type of research that can be done. Also the costs and the fact that the sample size is often even smaller than with a driving simulator pose limitations on what research questions can be addressed with instrumented vehicles.
- To overcome some of these disadvantages **naturalistic driving** was introduced. This method captures driver behaviour in a way that does not interfere with the various influences that govern those behaviours [6] and the (unobtrusive) observation methods takes place in a natural setting [7]. So no bias from observers, but there is a price in terms of experimental control. By not dictating routes, time slots, circumstances or other degrees of freedom there may be many confounding factors that make it difficult to draw very specific conclusions. The costs and usually still small sample size will also be major issues in the research design.
- A final option is to let road user answer questions about their behaviour at the end of trips or after a certain time period (revealed preference). These **self-reporting inquiries** could give better insight in decisions and motives, because they relate to real behaviour. The problem of socially desirable answers still exists, but could be reduced by the set-up of the questionnaire. However, it can be hard to find road users with vehicles equipped with the relevant C-ITS and/or automated driving functions.

What kind of data to collect is specified with indicators, which are used in hypotheses. A distinction can be made between measured indicators and estimated indicators. Examples of measured indicators related to traffic operations are: average local speed, total (cumulative) flow, (experienced) travel time, traffic composition, lane usage, travel choices or preferences and user compliance. Derived indicators are for example: travel time, density, space mean speed, total time spent, vehicle hours of delay, vehicle kilometres driven, user response to a measure, and costs and benefits. For evaluation of the driving task related to C-ITS and automated driving specific indicators for traffic operations and safety can be used [8].

An important step in the evaluation, which is often neglected, is selection of data. Which data can and should be used to verify the hypotheses? Normally this depends on the situational variables that could have influenced the experiment and/or the outcome of the experiment. Well-known factors that should be considered are weather conditions, occurrence of incidents, road works and other disturbances. But also time of year, traffic demand, route choice and other more operational traffic behaviour are important factors. It should also be considered if those factors are external or under investigation. In the first case they must be filtered out and in the second, on the contrary, they must be included in the selection.

3.4 Synthesis and conclusions

With the data collected the hypotheses can be tested. Valid and robust statistical tests are needed for this. For applying these tests two aspects are important:

- Is it allowed to use the test? For example: a t-test (Student's t) requires a Gaussian distribution of the samples, but this is often not checked and is also often not the case if real-life data are used.
- Certain tests require a certain sample size. In evaluating C-ITS the sample sizes are often small, so a check whether or not the test can performed is needed.

Most of the time it is difficult to determine significant effects, because variations (in traffic conditions as well as penetration rates) are large. But often it is still possible to draw meaningful conclusions. If significance is an issue, more research could be needed. But it is also necessary to think about when results are valid and how to determine that.

When all hypotheses have been tested, the research questions can be answered and conclusions can be drawn. It is good to keep in mind that the conclusions are only valid for the conditions in which the experiment was done.

3.5 Meta comparison

Results and conclusions from the experiment can be compared with the results of other experiments. This includes findings from literature and research done in other countries. The purpose of the meta comparison is to check and refine the validity of the results. Important aspects of the comparison are the definition of performance indicators, the experimental setup, measuring methods, calculations and analyses and the conclusions drawn. Often the same measures have been evaluated differently, which makes it difficult to compare the effects found (for example if the differences in travel time are measured for one measure, but not for the other). It is also important to check if differences in experimental setup and conditions are mentioned.

3.6 Generalisation

Arguably this step is the most important step of all. Generalisation means synthesizing and scaling the findings from previous steps such that they can be used (a) to support policy and decision making; (b) to focus and accelerate new R&D efforts, pilots and ultimately large-scale implementation of C-ITS and automated driving. Scaling means extrapolating results to higher penetration rates of technologies, extended time periods or larger geographic areas. This can be done in multiple ways. In some cases, that do not involve complex interactions between drivers or network effects, rules-of-thumb or simple relationships using data on situational variables can be used. In many other cases, the only method to address scaling issues is to use (combinations of) simulation models.

Given the validity problems of simulation models addressed above, it is of fundamental importance to understand these models as tools to test hypotheses; as "what if" engines, rather than precise representations of reality. Simulation models allow us to explore the consequences of the findings in pilot studies under higher penetration rates, in larger networks, etc. For example, what if drivers require a minimum of 10 seconds to safely take over the steering wheel after a period of 15 minutes of automated driving? What are the emerging traffic conditions under a series of different scenarios? What is the risk of accidents happening under many different vehicle compositions, circumstances, road-lay-outs, technologies?

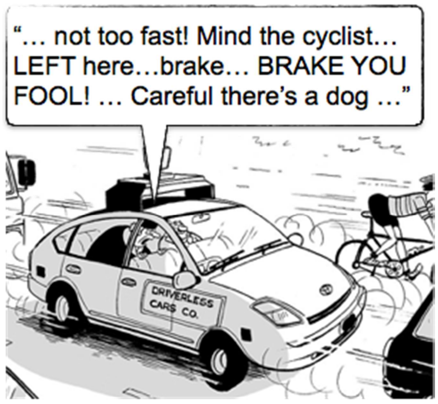
To answer these questions, we need to start using simulation models in different way. Instead of 10 or 15 replications, we need to start thinking in hundreds or thousands of replications where we systematically vary through Monte Carlo simulation over many different inputs and parameters. We may need to run these simulations in the cloud and not look at one single average outcome, but over distributions of possible outcomes. For example, C-ITS system X and Y may increase capacity between -10% to +30% depending on the circumstances (data and communication quality, traffic conditions, penetration rate, behavioural assumptions).

This notion of a range of possible outcomes empowers policy makers and to think in investment and development opportunities and risks and to develop policies for C-ITS and automation that are adaptive and flexible. Equally importantly, both the experimental results and the (possible) causal explanations that simulation models offer give the industry and road operators direction and focus for new R&D efforts, leading to a next cycle in the evaluation process.

4 Final remarks

We propose a cyclic approach for the R&D efforts in C-ITS and automated driving. This approach does not contain new evaluation methodologies—existing ones like FESTA can be used perfectly fine within this cycle. The main point is that a FESTA type field operational test *alone* is not sufficient. The evaluation cycle contains two additional steps in which the keyword is *generalisation*. How can we learn from each FOT and how can we generalise these lessons? What range of plausible and possible effects can we expect on efficiency and safety?

Even—or particularly—in the case that a pilot study does not yield positive results in terms of say road capacity, this continuous cycle of experimentation and generalisation accelerates the development of C-ITS. Failure leads to a much steeper learning curve than success. This is how science and technological invention work and this is in our view what it takes to succeed in large-scale uptake and implementation of C-ITS and automated driving. Lastly, a cyclic evaluation approach is highly cost efficient: the experimental designs, data processing tools and simulation models can be reused in a new cycle.



http://www.lawandai.com/2016/05/15/nhtsa-and-autonomous-vehicles-part-2/

Literature

[1] Vonk Noordegraaf, D., F. Faber en Th. Blondiau (2016). Inventarisatie van evaluatiemethoden voor C-ITS, memo voor TrafficQuest, 13 januari 2016.

[2] FOT-NET 2 Working Groups (2014). FESTA Handbook Version 5. Revised by FOT-Net (Field Operational Tests Networking and Methodology Promotion), 28 February 2014.

[3] Van Beek, P., S. van Lieshout en K. Malone (2016). De ontwikkeling van Connected ITS. NM Magazine, jaargang 11, nr. 2, pp. 9-13.

[4] Van Lint, H. (2016). Evaluatie is het vliegwiel van innovatie. NM Magazine, jaargang 11, nr. 2, pp. 15.

[5] Taale, H. (2016). Om het effect van C-ITS te bepalen, moeten we meer gebruik maken van modellen. NM Magazine, jaargang 11, nr. 2, pp. 14.

[6] Boyle, L. N., Lee, J. D., Neyens, D. M., McGehee, D. V., Hallmark, S., & Ward, N. J. (2009). SHRP2 S02 Integration of analysis methods and development of analysis plan, Phase 1 Report. University of Iowa, Iowa City, IA.

[7] Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., and Knipling, R.R. (2006). The 100-Car Naturalistic Driving Study, Phase II — Results of the 100-Car Field Experiment. Technical Report DOT HS 810 593, Virginia Tech Transportation Institute, Blacksburg, VA, April 2006.

[8] Harms, I. en Dicke, M. (2016). Rijtaakindicatoren voor C-ITS-projecten – Een verkenning naar een eenvoudig toe te passen set rijtaakindicatoren voor doorstroming en veiligheid. DITCM Innovations, 10 februari 2016.