UNIVERSITY OF TWENTE
CITO

Moral Reasoning in Traffic

About the possible relations between moral reasoning and traffic safety

N. Veldscholten
University of Twente

Prof. dr. ir. B.P. Veldkamp
(University of Twente)

Prof. dr. ir. T.J.H.M. Eggen
(University of Twente)

Dr. E.C. Roelofs
(Cito)

June 2015

Table of Contents

Foreword / Acknowledgement

I would like to use this space to thank a few parties that were involved in the process of writing my final thesis.

Writing this thesis would not have been possible without the support and guidance of Cito, Arnhem. By joining the POK-department (Psychometrisch Onderzoeks- en Kenniscentrum) as an intern, I had optimal access to all relevant facilities and information I needed for this study. Thanks to this department, I met a lot of people during my internship who inspired me when writing this thesis.

Thanks to everyone behind the DriveXperience, the program that made it possible for me to include this significant number of respondents in this study.

Special thanks to Erik Roelofs, who offered me the opportunity to work on this research, and who provided me with information and guidance from beginning to end.

Nienke Veldscholten
June 2015

Summary / Abstract

In choosing to participate in traffic in a responsible manner, not only cognitive skills are needed, but also socio-moral reflection. Three assessments were developed to measure the level of moral reasoning in traffic situations. Also, possible relations between moral reasoning and traffic safety were studied, including variables as the involvement in accidents, driving behaviour, motives to comply with the rules, mileage and personal characteristics as age and gender. The results indicate that the lower levels of moral reasoning are related to a higher number of accidents, to a higher driving speed, and to a higher degree of space-taking behaviour. Furthermore, the findings suggest that men show more risky driving behaviour than women, as well as younger drivers do than older ones. With reference to these conclusions, it might be useful to pay more attention to moral behaviour in traffic in second phase driver's programs. This might contribute to lower accident rates in the target group of young drivers.

Introduction

Second phase driver training program

Since 2003 a so-called second phase driver training program is offered in three Dutch provinces, namely Gelderland, Zuid-Holland and Limburg: the DriveXperience (Roelofs, Vissers & Van Onna, 2013). The goal of this program is to reduce the risk of accidents among young drivers who have just passed their driving exam. In order to achieve this reduction, the program addresses critical self-reflection on one's own driving behaviour and the development of self-regulation skills.

Higher order skills such as critical self-reflection and self-regulation skills have become more important in driver training. This emphasis originated in studies on the effects of driver training, which proved repeatedly that learning technical skills and knowledge of traffic rules and signs did not have an effect on driver safety, and sometimes even reversed effects (Christie, 2001; Ker et al., 2003). Second phase driver training should therefore emphasize the way young people choose to drive (Elander, West & French, 1993; Helman, Grayson & Parkes, 2010).

Because driving is a social activity that takes place in a wider context in which the driver constantly interacts with its immediate environment and the vehicle (Rakotonirainy & Maire, 2005), the presumption was created that moral reasoning in traffic and the reflection on it might be relevant factors in competent driving. In choosing to participate in traffic in a responsible manner, not only cognitive skills are needed, but also social-moral reflection. The basic idea behind the presumption is that the moral development level of young drivers will be reflected in their ability and willingness to empathize with the interests and needs of others.

Problem statement: Self-assessments and moral reasoning

An integral part of the DriveXperience is the use of self-assessment among the participants, which yields feedback reports about the driving style, which are expected to provoke self-reflection on the part of the young driver. The assessments developed in this context are provided by CitoDrive, a part of Cito aimed at the training of license holders which involves formative assessment with a diagnostic function (Roelofs & Vissers, 2010). As a preparation for the Coaching Day, young drivers complete the online Driver Risk Assessment (DRA) and the Situational Awareness Test (SAT). They receive a personal report about hazard recognition, awareness of traffic, driving style, and risk profile. Subsequently, they go through the Driver Self-Assessment (DSA), which provides insight into the driver's self-image. By these results, the driving coach knows where to pay attention to.

In the present study, three additional online self-assessments were developed, which aim to measure the level of moral reasoning in traffic situations. In addition, the reports should elicit socio-moral reflection on the part of the young driver. On top of that, driver coaches can make recommendations regarding the attitude towards traffic rules and other road users while coaching a young driver.

The aim of the study is to develop valid and reliable assessments to measure the level of moral reasoning in traffic situations, and to test whether these assessments are able to measure possible relations between variables that could influence the moral reasoning of young drivers.

Conceptual Framework

Moral reasoning

A definition of moral reasoning was derived from Treviño, Weaver and Reynolds (2006) by Reynolds and Ceranic (2007), who describe moral and ethical behaviour as "behaviour that is subject to (or judged according to) generally accepted moral norms of behaviour". They add that moral behaviour occurs within the context of larger social prescriptions. Moral behaviour is influenced by moral identity and moral judgment. When someone perceives himself as a moral individual, he will act in a more moral manner and will pursue more extreme alternatives (always or never showing certain behaviour) than someone who does not see himself as moral (Reynolds & Ceranic, 2007; Colby & Damon, 1992). Moral judgment includes what someone perceives as right or wrong (Kohlberg, 1983).

The way in which people engage in moral reasoning depends on their moral development. Moral development contains the development of someone's conscience, in other words the awareness of what is right and wrong (Kuhlemeier, Boxtel & Van Til, 2011). This means that someone is able to indicate why exactly something is morally right or wrong (Beerthuizen, 2012). Earlier research proved that a high level of moral reasoning is associated with less aggressive behaviour, less delinquent behaviour, and more pro-social behaviour (Blasi, 1980; Stams et al., 2006).

Gibbs, Basinger and Fuller (1992) distinguished four stages of moral development. In the first two stages (the immature level) moral judgment development is characterised as immature and superficial. Stage one is about centrations on or over-attention to salient features such as size and power, whereas stage two contains pragmatic or instrumental exchanges or concrete moral reciprocity. The third and fourth stage (the mature level) are characterised as mature and profound (Gibbs, Basinger & Fuller, 1992; Kuhlemeier et al., 2011). In the third stage judgment contains mutualities or ideal moral reciprocity, and the fourth stage is about systems; the expansion of mutualities into complex social systems. The stage 3 mutualities would be sufficient for 'the traditional and isolated peasant village', but this familiar way of communication may not need the 'more formal and elaborate legal mechanisms' and standards that would be necessary for stage 4 contexts, such as dispute resolution and social equilibrium in a more heterogeneous or pluralistic complex society (Gibbs, Basinger, Grime & Snarey, 2007). These stages are illustrated in Table 1.

Table 1. Four stages of moral development (Gibbs, Basinger & Fuller, 1992)

	Level A	
Stage 1		Centrations: physical and short-term
Stage 2		Pragmatic exchange
	Level B	
Stage 3		Reciprocity
Stage 4		Systems

Moral development consists of different aspects. Gibbs et al. (1992) addressed the following aspects of moral development: justification, empathy and cognitive distortions.

Moral justification of made choices is an aspect of moral behaviour. It refers to the moral awareness of consequences a certain action could have. The justification and the ability to apply moral criteria to evaluate behaviour is crucial in this context. It pertains to the ability to reason why something is morally right or wrong. In Table 2 the aspect justification is divided into the stages of Gibbs et al. (1992). It can be seen that the lower, immature stages are about extrinsic motivation, wanting to prevent punishment or other problems, about pragmatic needs and one's own preferences. Justification on the mature stages can be described as pro-social, it is intrinsically motivated, reciprocity and the social functioning are central, and someone on this level values his or her reputation and integrity. A higher level of moral justifying is related to less aggressive, delinquent and antisocial driving behaviour (Blasi, 1980; Stams et al., 2006). When someone justifies antisocial behaviour as something negative, it diminishes the chance that they will display the same behaviour (Kuhlemeier et al., 2011). Earlier research indicates that the attitude towards antisocial behaviour of risk students is a predictor of the later delinquency (Leenders & Brugman, 2005; Nas, Brugman &

Koops, 2005; Vitaro, Brendgen & Tremblay, 2000).

The second aspect, empathy, is an important part of social competence and moral development (Hoffman, 2000; Ten Dam & Volman, 2000). Empathy means that someone is able to recognize emotions, to take someone else's emotional perspective and to sympathize with another (Berk, 2006). Applied on driving, empathy means that driver is able to empathize with the thoughts, emotions and intentions of other road users. This also means that the driver shows empathy himself, including feelings such as concern, sympathy and compassion (Kuhlemeier et al., 2011). Young people with severe behavioural problems have less empathy than young people without problems (American Psychiatric Association, 1994; Stegge, 2006). The different moral stages of empathy are illustrated in Table 2. The first stage contains self-interest and not taking account for the interest of others; someone drives the way he wants. In the second stage the driver is aware of others, but in cases of impediments or annoyances he does not forgive other road users. In the higher stages there is more attention to other road users: their well-being and the public interest is placed before the individual interest. In the highest stage the driver knows why something works and why it is important to act in an empathic way.

Table 2. Aspects of moral reasoning according to the stages of Gibbs (1992)

	Moral reasoning	
Stages	**Justification**	**Empathy**
Stage 1 *Physical and short-term*	- Preventing punishment (fine, warning) - Preventing problems with authority (law, rules, person)	- Self-interest is central - Does not account for the interest of others
Stage 2 *Pragmatic exchange*	- Own preferences are central - Pragmatic needs - Eye for an eye-attitude - Is extrinsic motivated	- Is aware of others - Is capable of taking the perspective of others - Is not able to 'forgive' others
Stage 3 *Reciprocity*	- Pro-social-attitude - Is aware of the consequences of the own reasoning on others - Reciprocity - Is intrinsic motivated	- Takes a more empathetic role - Pays attention to the well-being of others
Stage 4 *Systems*	- Social functioning is central - Values reputation and integrity	- The public interest is placed for the individual interest - Knows why something works, why something is significant

A lagging moral development does not automatically result in severe behavioural problems. According to Gibbs (2003), self-serving cognitive distortions are needed to get those behavioural problems. Cognitive distortions are incorrect or biased beliefs, assumptions or interpretations about the social behaviour of someone else or yourself (Barriga, Gibbs, Potter & Liau, 2001; Gibbs, 2003). According to Gibbs (2003), everybody is susceptible for cognitive distortions. The more often they occur, the bigger the chance that a lagging moral development results in the occurrence of behavioural problems (Gibbs, 2003). There are two types of cognitive distortions: primary and secondary. Primary distortions are about having a strong selfish thinking perspective, which means that someone does not take into account the feelings, wishes, needs and rights of others. When having secondary distortions, the antisocial behaviour is rationalised. This means blaming others, minimizing the effect of someone's own behaviour and assuming the worst (Brugman et al., 2011; Gibbs, 2003; Kuhlemeier et al., 2011). These cognitive distortions are considered self-protective, since they serve someone's self-image. By employing these distortions, problem behaviour becomes acceptable for this individual and they help ensure that the positive self-image is not getting harmed. The distortions offer a justification for the behaviour and minimize feelings of shame and guilt (Kuhlemeier et al., 2011). The different

moral stages are not applicable for cognitive distortions, since they exclusively take place at lower moral levels. Because of that, only this aspect is not divided in stages.

Driving as moral task domain

In order to relate moral behaviour to young drivers driving, it is important to define the task domain of driving. Driving is a complex task, partly due to the dynamic environment, which can cause dangerous situations since it requires continuous attention (Van Emmerik, 2004; Kettwich, Stocky & Lemmer, 2010). Driving can be seen as a lifelong learning process, since driving can be trained and developed continuously (Vissers, Van Betuw, Nägele, Kooistra & Harteveld, 2007; Van Emmerik, 2004). Nowadays, the focus of driving schools is to let their student drivers drive more social.

In driving courses four levels should be addressed to coach a student driver adequately, according to the Goals of Driver Education (GDE)-matrix of Siegrist (1999). This matrix is shown in Table 3.

Table 3. Goals of Driver Education (GDE)-matrix (to Siegrist, 1999)

Level of behaviour	Knowledge and skill	Risk-increasing aspects	Self-assessment
1. Goals for life and skills for living	Lifestyle, age, group, culture, social position, etc. vs. driving behaviour	Sensation seeking, risk acceptance, group norms, peer pressure	Introspective competence, own preconditions, impulse control
2. Goals and context of driving	Modal choice, choice of time, role of motives, route planning	Alcohol, fatigue, low friction, rush hours, young passengers	Own motives influencing choices, self-critical thinking
3. Driving in traffic	Traffic rules, cooperation, hazard, perception, automatization	Disobeying rules, close-following, low friction, vulnerable	Calibration of driving skills, own driving style
4. Vehicle control	Car functioning, protection systems, vehicle control, physical laws	No seatbelts, breakdown of vehicle systems, worn-out tyres	Calibration of car-control skills

Most driving courses only address the knowledge, skills and risk increasing aspects at the third level (driving in traffic) and the fourth (vehicle control) within the grey area. This means self-assessment is normally not included, as well as all parts belonging to the first level (goals for life and skills for living) and level 2 (goals and context of driving) (Deery, 1999; Vlakveld, 2006). According to this, moral reasoning is in general not (enough) addressed in driving courses. Despite the necessity of addressing this issue, articles that concern driving skills are not often about moral driving skills. This makes moral reasoning in traffic a relatively unexplored topic.

Moral reasoning in traffic or driving socially means that a driver not only pays attention to his or her own safety and traffic flow, but also to the needs of others. This means that he accepts the mistakes of others and not only sticks to his own rights. Also, he tries to prevent surprising or hindering others. A way to accomplish this, is to make driving intentions clear in time. Social driving also means that the driver allots enough space to others to let them be able to correct errors. Furthermore, he ignores the aggressiveness of others and he is able to sympathize with other road users (Roelofs, 2013).

There are some variables of influence in social driving. It has been proven that demographic variables as age, gender and driving experience are relevant to predict antisocial or aggressive driving behaviour (Moore & Dahlen, 2008). Also forgiving and the consideration of consequences contribute to the prediction of aggressive driving and the expression of anger. Next to these indicators, there are a few more personality factors that could implicate an aggressive driving style, such as looking for sensation, impulsiveness, and the characteristics of driving anger (Dahlen, Martin, Ragan & Kuhlman, 2005; Jonah, 1997; Schwebel, Severson, Ball & Rizzo, 2007). A driver who tends to consider the consequences of his behaviour drives less aggressive, since the consequences of aggressive driving are more striking than those of driving in a more moral way.

Another variable that may be of influence during driving in traffic, is the driver's self image. Research showed that most drivers tend to believe that they are better drivers than the average driver (Näätänen & Summala, 1975; Svenson, 1981). Svenson (1981) showed that 88% of his American respondents believed themselves to be safer than the average driver. This study suggests that this may be mainly a result of lacking information about other drivers, which may lead drivers to regard themselves as "better". Svenson (1981) stated that this (unjustified) believe may lead to greater risk in driving, which is positively reinforced for those who actually are successful in driving.

Young drivers. Young novice drivers are overrepresented in road accidents (Deery, 1999). A quarter of the drivers between 18 and 24 years old gets involved in a traffic accident within four years after passing their driving exam (Roelofs & Vissers, 2010). In addition, Ryan, Legge and Rosman (1999) studied age and gender differences in the rates of crash involvement, and found that drivers younger than 25 years were involved in 35% of all crashes. When looking at crashes in daylight, the percentage of crashes for drivers under 20 years turned out to be 64%.

Yagil (1999) found that younger drivers and male drivers express a lower level of motivation to comply with traffic laws than female and older drivers, and the results of the study of Al-Balbissi (2003) showed that male accident rates are significantly higher than female ones. Explanatory variables for these accident rates are lack of attention and impatience among male drivers.

The reason that young drivers tend to be more often involved in accidents, could be explained by the fact that they get involved in tasks they cannot handle yet (De Craen, 2010). Therefore it is important for drivers to choose traffic situations they can manage. When the task requirements of a traffic situation are higher than a driver's level of driving skills, he will lose his control of the situation (Fuller, 2005). This can result in an accident, a compensating action of another road user or a lucky escape.

Driving as a domain of competence. Drivers differ in their degree of competence development. A skilled driver reflects regularly on his decisions and actions. Reflection leads eventually to the adjustment of someone's own knowledge, skills, attitudes, emotions, and mood (Roelofs & Sanders, 2007). This reflectively, cyclical process is shown in Figure 1.

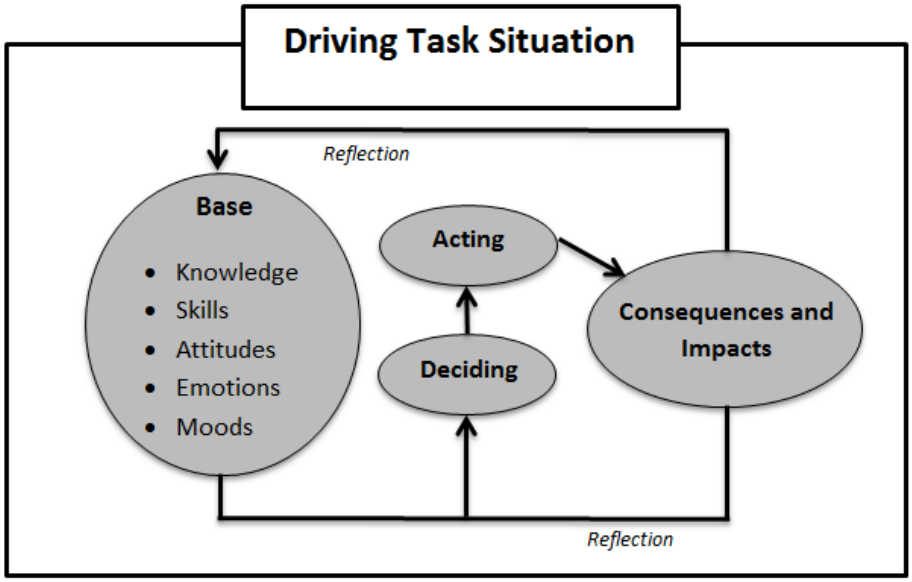


Figure 1. Reflectivity within a driving task situation

Driving is more than just an individual activity. Driving involves a combination of some mental task processes, which should be addressed during the driver training together with their effects (Roelofs, 2013). During driving, there is a continuously cognitive-affective decision process, which includes a chain of mental processes: perceiving, predicting, empathizing with others, evaluating, deciding, and acting. While going through these processes, a driver has to take into account his own situation, the situation of possible other passengers, other road users, heavy traffic, weather conditions, etcetera. In

addition, it may occur that the driver performs different competing actions while driving, such as phoning, texting, talking, irritating to others, and so forth.

In Figure 2 the moral decision model is shown. It is specified which mental actions the driver performs during driving. This model can be applied on every road user. It is a cyclic process that usually starts with perceiving other road users and empathizing with them. Then the driver becomes aware of the moral issue and makes a moral consideration, after which he decides what an appropriate judgment is in the concerning situation. The driver establishes the intention to act morally and eventually he engages in moral behaviour (Roelofs, 2013; Reynolds & Ceranic, 2007). The action that follows can have a direct impact on the traffic situation, but the driver self can also experience a direct impact. This way, he reflects on his own decision process (Roelofs, 2013).

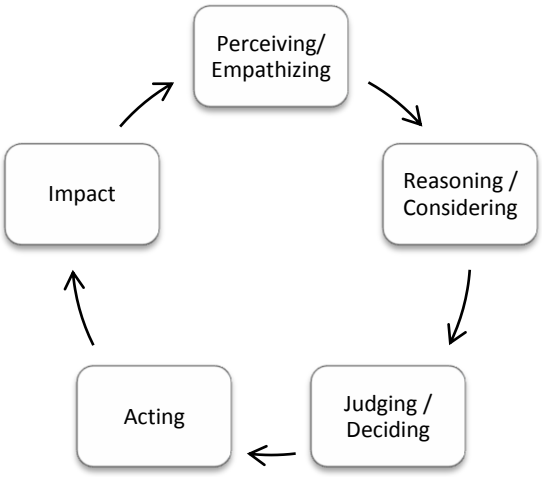


Figure 2. Moral decision model

It can be concluded that driving as a complex task with its mental and decision processes requires much mental task processes. To have control of a situation, a driver's skills need to be developed enough to exceed the level of task requirements. The more skilled the driver, the more he reflects on his driving behaviour. The more a driver reflects, the more he probably tends to engage in moral behaviour.

It can be expected that young drivers with a low level of moral reasoning show more often traffic violations and aggressive driving behaviour, than young drivers with a higher level of moral reasoning (Bianchi & Summala, 2002).

Assessment of moral reasoning

To establish one's moral level, this behaviour needs to be measured by means of an assessment. Black and Wiliam (1998) define assessment broadly to include all activities that teachers and students undertake to get information that can be used diagnostically to alter teaching and learning. According to this study, assessments include observations, discussions, and analyses of one's work.

Because the emphasis of driving is changing to higher order skills and the training of it becomes more competency-based, new ways of driving assessment are needed (Roelofs, Vissers & Van Onna, 2013). The new opinions about driving are nowadays included in learning goals for driving training (Hatakka, Keskinen, Gregersen, Glad & Hernetkoski, 2002). A taxonomy of learning goals is established, where the higher order skills are distinguished, such as risk perception and reflection on the own driving behaviour.

Driving behaviour can be assessed by both formative and summative assessments. A summative assessment is a judgment, which contains all evidence to a certain moment of what a student knows and does not know (Taras, 2005; Garrison & Enhringhaus, 2007). This moment is seen as a finality at the point of the judgment. A summative assessments and its functions do not influence the process (Taras, 2005), and can only help in evaluating certain aspects of the learning process, such as effectiveness of programs. It takes a formative assessment to provide information and to make instructional adjustments and interventions during the learning progress (Garrison & Enhringhaus,

2007). Formative assessment is pedagogic and cannot be separated from instruction (Garrison & Enhringhaus, 2007; Black & Wiliam, 1998). It is not only about collecting information or data on student learning, but also about what teachers or trainers do with the gained information. According to Black and Wiliam (1998), an assessment only becomes formative when the information is used to adjust teaching and learning to meet the needs of students. In this case: to develop the moral reasoning of the young drivers.

Furthermore, practice is formative when, after eliciting, interpreting, and using the information, decisions are made about the next steps in instruction that are likely to be better, of better founded, than the decisions that would be taken in the absence of the information that was elicited (Black & Wiliam, 2009). The concerning instruction contains any activity that is intended to create learning, and the decisions are actions to require that the elicited information is used to make instructional adjustments that improve learning. This will not always result in better learning for all students.

Formative assessment makes it thus possible to use the gained information for learning purposes, since the learning goal is central (Gipps, 1994). A formative assessment requires feedback that indicates the existence of a 'gap' between the actual level of the work being assessed and the required standard, and an indication of how the work can be improved to reach the required standard (Taras, 2005). Useful feedback is thus considered to be a critical component of formative assessment (Kingston & Nash, 2011). By providing students with descriptive feedback as they learn, they will engage in the assessment of their own learning (Garrison & Enhringhaus, 2007). This kind of feedback is the most significant instructional strategy to stimulate students in their learning. It provides them with an understanding of what is going well, and it gives them specific input on how to continue to the next step in the learning progress. According to Nicol and Macfarlane-Dick (2006) good feedback practice strengthens the students' capacity to self-regulate their own performance. They suggest seven principles about good feedback:

1. It helps clarify what good performance is (in terms of goals, criteria, expected standards)
2. It facilitates the development of self-assessment (reflection) in learning
3. It delivers high quality information to students about their learning
4. It encourages teacher and peer dialogue around learning
5. It encourages positive motivational beliefs and self-esteem
6. It provides opportunities to close the gap between current and desired performance
7. It provides information to teachers that can be used to help shape the teaching

In addition, to contribute to the learning process of the driver, it is of great importance that the formative feedback is related to aspects of the driving task and situation, that they match the stage of driving development of the driver, and that they provide information about the mental processes that occur (Kuiper, De Penning & Roelofs, 2011).

To establish the driving ability of the driver during the training, these formative assessments can be conducted as practical assessments, (digital) surveys and self-evaluations (Roelofs & Vissers, 2010). It has been demonstrated that web-based methods of collecting data yield results comparable to those of traditional formats, such as paper and pencil (Krantz & Dalal, 2000).

Research questions

There is a need for an instrument about moral reasoning. The main purpose of these assessments is to create a profile for the young drivers about their reasoning in traffic before they start with their driving course. With the help of this profile, the young driver and driving coach gain insight into characteristics which could be of influence on the driving style. Risky attitudes, opinions, moral choices and behavioural intentions could be identified this way. The focus of the assessments is to gain an understanding about the level of moral reasoning in traffic of the young driver. The conceptual framework is portrayed in Figure 3.

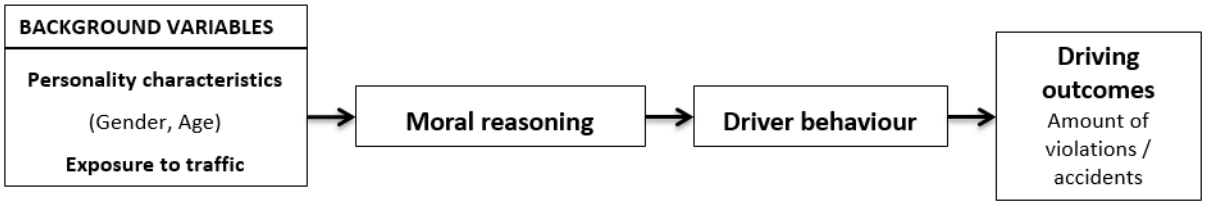


Figure 3. Model of conceptual framework

The main question of the present study is: To what extent are the developed assessments about moral reasoning reliable and valid?

In this context, three specific research were studied:

1. To what extent are the developed subscales for moral reasoning reliable and valid?

2. Which levels of moral reasoning are perceived among young drivers of the DriveXperience?
    a. To what extent do young drivers justify their norm-complying behaviour according to the different levels of moral reasoning?
    b. To what extent do young drivers justify possible antisocial driving behaviour?
        i. To what extent do young drivers use cognitive distortions to justify antisocial driving behaviour?
        ii. To what extent do young drivers use specific types of cognitive distortions?
    c. To what extent do young drivers allot space to other drivers? And: to which degree do young drivers take space at the expense of other drivers?

3. To what extent is moral reasoning related to driving behaviour, such as amount of violations and accident involvement? To what extent is there a relation with personality characteristics such as gender, and age?

Hypotheses

Based on the theory, possible relations between the variables are indicated. With reference to these expectations, the following hypotheses are formulated:

a) Lower levels of moral reasoning are related to a higher number of accidents
b) Lower levels of moral reasoning are related to a higher driving speed
c) Lower levels of moral reasoning are related to a higher degree of space-taking behaviour
d) Lower levels of moral reasoning are more related to male drivers than to female drivers
e) Lower levels of moral reasoning are more related to younger drivers than to older drivers

Study design

Assessment design

For the development of the assessment to measure moral reasoning, the *evidence-centered design* (ECD) of Mislevy, Steinberg, Almond and Lukas (2006) is used. The ECD-model is developed to assess interactive tasks, multidimensional abilities, and complex performances. Moral behaviour during driving can be perceived as a high-performance task (Van Emmerik, 2004), which makes the ECD-model an adequate design model.

Mislevy et al. (2006) describe how the development of valid assessments requires explication of the different sub-models. By using the ECD-model, it is possible to set up validity arguments regarding judgments and decisions about students (Kane, 1992; Kane, 2006). The validity argument starts with the description of the knowledge, skills or other attributes that should be assessed (Mislevy & Haertel, 2006). Also, it should be described which behaviour or performance the competence entails, and in which situations this behaviour will be elicited. Mislevy and Haertel (2006) distinguish five interactive layers of assessment design: the domain analysis, the domain modelling, the conceptual assessment framework, the assessment implementation, and the assessment delivery. These layers are shown in Figure 4.

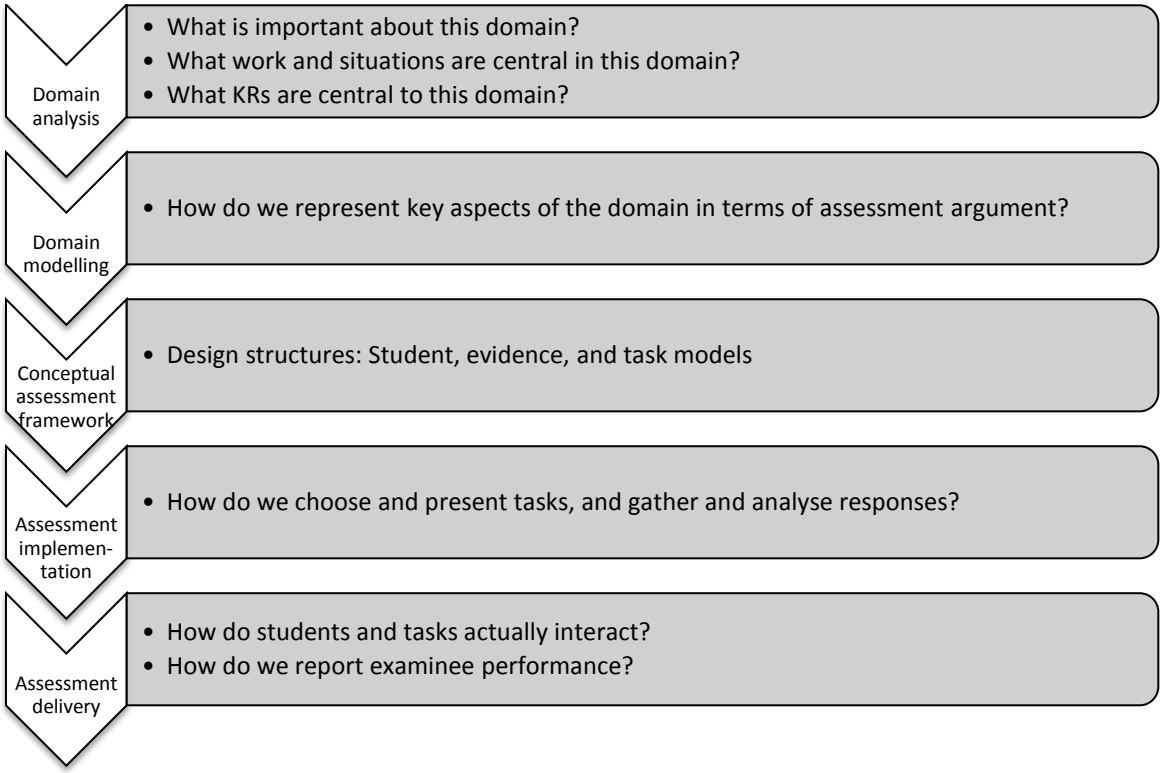


Figure 4. Evidence Centered Assessment Design

The domain analysis and modelling have been discussed in the theory section, and the assessment implementation and delivery are not explicitly relevant in the context of the involved research questions and will thus not be discussed. The emphasis of this research is on the third layer: the conceptual assessment framework, which contains the structure of the final design of the assessment. This could include the form of the structure and specifications for tasks and assessment, evaluation procedures, and measurement models.

The conceptual assessment framework is formed by three sub-models: the student model, evidence model and task model. Next to those there are two other models: the assembly model, and the presentation model. The five sub-models together describe parts that need to be addressed for the development of the assessment. Roelofs (in preparation) added a sixth model: the report model, which is also included in this study. The value of this additional model can be explained by the fact that reports provide guidance to adjust training and to adapt it to the needs of individual students. Not

included in the assessment design are the assembly model and delivery model. Those are about how the assignments are selected, how they are delivered and when the delivery is done, which is already covered in the presentation model. In Figure 5 the models for assessment design are presented.

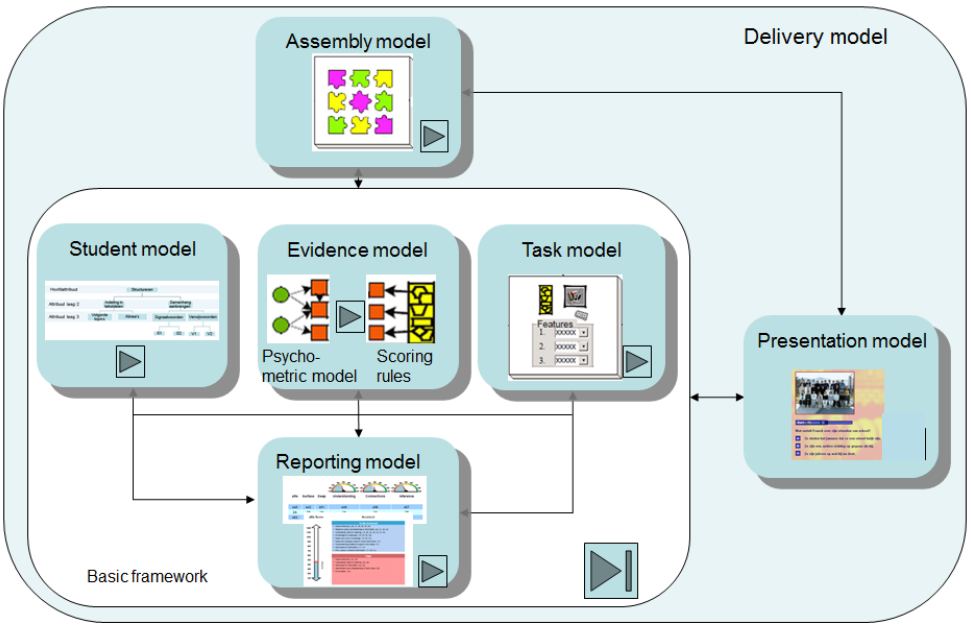


Figure 5. The ECD-Assessment model

Student model. The student model describes the nature of the skill that is going to be assessed. It is summarized which mental processes are involved when the student performs certain tasks within the domain or field. The model provides a structured overview of the attributes, which include the essential knowledge, skills, strategies, and attitude elements that need to be reported by the assessment. It is described in which task situations the attributes should be used, which learning progression they go through, and which barriers or obstructions have to be overcome. Often a rearrangement or reformulation of the goals and sections is needed until there is a logical structure of attributes, to make the assessment work in practice.

The three (web-based) self-assessments are based on Gibbs' theory (Gibbs, 1991) about moral development. The first is a self-evaluation, which measures the extent to which respondents use cognitive distortions, in other words the reasons to justify non-social or violating behaviour; the second is a self-assessment about eighteen common traffic situations with a possible space conflict, in which the respondent can choose to take space at the expense of another road user, or to share or allot the space to others, as a measure of empathizing with other road users; and the third assessment is a survey about fifteen traffic situations, in which the respondent is asked how many times he commits a proposed traffic offense, and the reasons are asked why someone would or would not commit the offense, as a measure of levels of justification.

Task model. The task model describes a) the key characteristics of the tasks and the task situation, which are needed to make it possible to observe the thoughts and actions of the students; b) the way in which the tasks are offered, such as instructions, the structure, stimuli, and prompts; c) the expected responses of the students, such as actions or products. This model can be seen as the blue print of the assignments.

In all these traffic situation items, a lot of characteristics can be of influence. In this study, two main aspects of critical influence are mostly taken into account: the presence of other road users at the same time, and the 'action space'. In making moral decisions, it is a condition that other road users are present. This may include road users who arrive at the same time at an equivalent crossing, or at multiple crossings with more different traffic users. The driver has to oversee the situation and make a decision in a relatively short time. The space in which a driver has to make decisions and act appropriately, increases the difficulty of the traffic situation. Both the action space as the action time

are influenced by the driven speed. For example, a road narrowing may influence this action space (Roelofs, 2013).

Evidence model. The evidence model is the bridge between the student model and the task model. It ensures that the obtained answers are converted to scores on main attributes, part attributes and sub-attributes. The model consists of two parts: the first is the scoring model, which includes rules about assigning scores to the results of the students. The second part is the measurement model or psychometrical model, wherein the perceived measures are linked to the presumed attributes in the student model and the relations in-between. This part also provides measurement scales, by which statements can be made about the attributes of the students, such as their skills and potential underlying processes.

For the first assessment about cognitive distortions the score indicated to which degree the respondent used cognitive distortions to justify antisocial behaviour, and to what extent the four different categories are used. This was scored with a 6-point Likert-scale. For the second and third assessment, the marble method was used. Hundred point had to be divided over three (second assessment) or four (third assessment) options. This is based on the procedure of Nedelsky (1954), which forces the rater to give detailed consideration to the specific content of an assessment, rather than to its general characteristics. The technique supposes that a student proceeds by eliminating incorrect alternatives and then chooses at random between the remaining alternatives (Van der Linden, 1982). This way, for the second assessment about empathy it was possible to indicate if someone used an egocentric or altruistic perspective. For the third assessment about justification, the marble method made it possible to indicate which motive category to justify norm-complying behaviour was used most frequently by a respondent.

Report model. Roelofs (in preparation) added this model, since the information that has to be reported has great influence on the choices that have to be made for the construction of an assessment. The most important question in context of reporting is the granularity of reports. It has to be decided if the report is going to discuss a skill in general, or more specific by describing part attributes of the skill. A more fine-grained report also means a more fine-grained student model and task model.

The respondents received three feedback reports, one for each assessment. It was chosen to show their results in graphs instead of text. An example derived from the assessment about cognitive distortions can be seen in Figure 6. The used colours give an immediate impression of good or bad results. All assessment scales vary from zero to hundred, divided in subscales of 25 points: very unlikely, unlikely, likely, very likely. The first two of these scales are (shades of) green, which indicate a good score. The last two are (shades of) red, which indicate a bad score. In that case it would mean that the respondent would (very) likely use cognitive distortions in traffic. The score of 28 in the example means that the chance that the respondent uses cognitive distortions in traffic is unlikely.

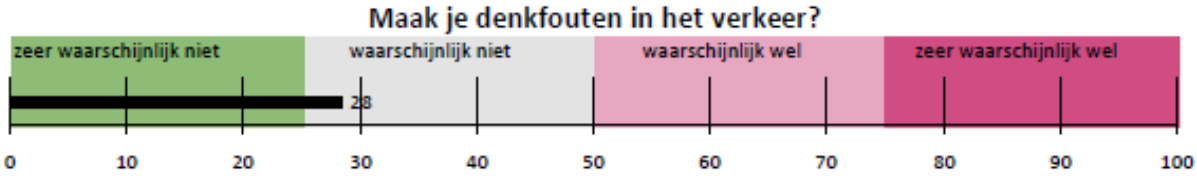


Figure 6. Report example Assessment 1: Cognitive Distortions

In Figure 7 an example of the report for the second assessment about empathy is illustrated. Again, the colours indicate the cut-off scores. A (red) score lower than 40 means an egocentric perspective (takes space at the expense of others in traffic), a (green) score higher than 80 means an altruistic perspective (allots space to others in traffic). The score of 77 in the example means that the respondents allots space to others (a bit) more often than taking the space.

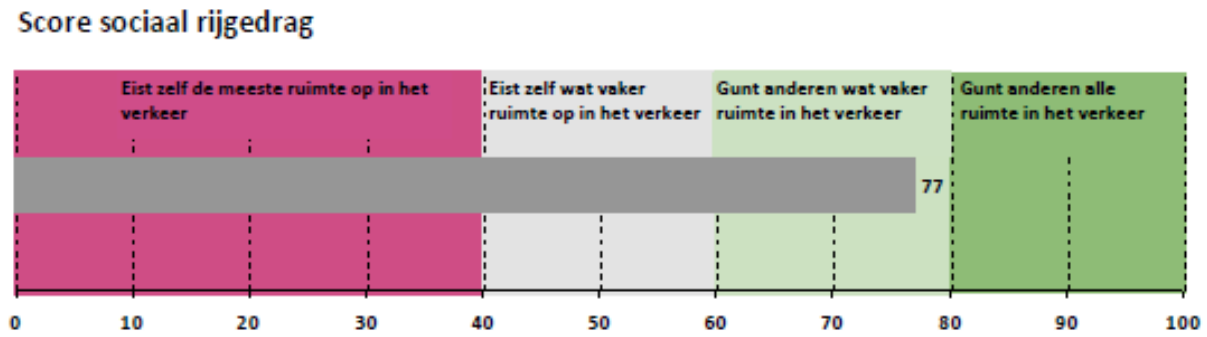


Figure 7. Report example Assessment 2: Empathy

In Figure 8 an example of the report for the third assessment about justification is shown. A (green) score lower than 15 means that the respondents show no norm-exceeding driving behaviour, based on the results of the assessment. A (red) score higher than 60 means that someone shows very regularly norm-exceeding behaviour. The example shows a score of 15, which is between 'no' and 'limited' norm-exceeding behaviour.

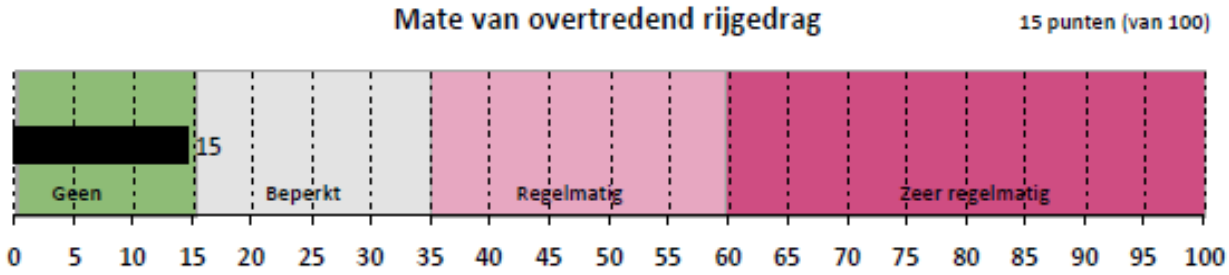


Figure 8. Report example Assessment 3: Justification

A more detailed explanation of the feedback report is given in the chapters about the developed instruments.

Presentation model. In the presentation model it is described how the assignments are presented to the students, in both the physical and social situation in which the tasks occur. This contains all technical, material and social conditions that are needed to evoke task performance, which can be achieved in three ways: written, digital, or in a (simulated) live or work situation. The first three models (student, task, evidence) are guiding for this model. Restrictions in the presentation model could lead to adjustment in the assessment design, and thus also in the contemplated conclusions about students. For instance, it could turn out that the desired stimuli or response methods are too expensive or too difficult to use in the available technical infrastructure.

The assessments are presented in SurveyMonkey, a web-based tool for developing and sending out questionnaires. Respondents are easier to reach this way than in person, and furthermore, online surveys offer more possibilities than traditional paper surveys (Baarda, De Goede & Teunissen, 2009). Besides, a survey is the most suitable design to get an overview of opinions of the respondents under most natural conditions (Baarda et al., 2009).

For the assessments about empathy and justification, traffic situations were presented as items. In the justification assessment, chalk drawings were used instead of real images as can be seen in Figure 9. This way it was expected that the respondent would be able to see the situation more as a hypothetical one rather than a recognizable one. Consequently, this would result in more representative results for similar situations. For the empathy assessment this measure was less required, since those items are more specific and also had to evoke more empathy than the justification items about general violations.

Figure 9. Example of 'chalk drawing'-item

Developed instruments

Assessment 1: Traffic situations. In Appendix 1 a print screen of Assessment 1 can be found. The first assessment includes the aspect of cognitive distortions. An existing questionnaire was used to formulate and develop new items: the questionnaire How I Think (HIT) (Barriga et al., 2001; Brugman et al., 2011; Kuhlemeier et al., 2011). This assessment is based upon the four categories of self-serving cognitive distortions of Gibbs and Potter (Gibbs, 1991, 1993; Gibbs, Potter & Goldstein, 1995). These categories are also used in this study:

- Self-centered
- Blaming others
- Minimizing / Mislabelling
- Assuming the worst

The assessment consists of 38 items of cognitive distortions, which are derived from common violations and annoyances in traffic (Infonu, 2014), and from the four-category HIT-questionnaire. The category 'Self-Centered', a primary cognitive distortion, is about the extent to which someone prevails his or her own needs, interests and feelings above those of others (e.g. *What I consider important in traffic, is that I can drive how I want to drive*). The other categories are secondary cognitive distortions. The second category, 'Blaming Others', concerns the extent to which someone blames others for his or her own antisocial or aggressive driving style (e.g. *If I drive close behind another, it is because the other is not fast enough*). The third category, 'Minimizing / Mislabelling', includes items where the driver thinks that violations are acceptable or will barely/not do any harm to others (e.g. *If others don't follow the rules, then I also don't have to follow them*). The last category, 'Assuming the Worst', is about the extent to which someone wrongly attributes hostile intentions to others (e.g. *Other drivers intentionally bother me*).

When taking into account the moral decision model (Figure 2), the next steps are included in this assessment: after reading the item, the respondent considers his options (*reasoning/considering*), than he decides what he would do in a similar situation (*judging/deciding*) and which action would follow (*acting*). This will result in a certain effect, depending on the decisions and acts of the respondent.

The items contain different violations and types of anti-social behaviour, as shown in Table 4. These are based on frequently occurring violations. The categories of 'speed violation' and 'short following distance' are most included in the assessment (respectively $k = 7$ and $k = 5$).

Table 4. Violations and anti-social behaviour

Act of driving behaviour	Frequency
Speed violation	7
Short following distance	5
Victim role	4
Violating the rules	3
Justifying your acts	2
Losing patience	2
Merging (right to left)	2
Overtaking	2
Parking	2
(not) giving priority	1
Cutting off	1
Honking the horn	1
Indicating direction	1
Merging (left to right)	1
Not moving	1
Phoning while driving	1
Traffic light	1

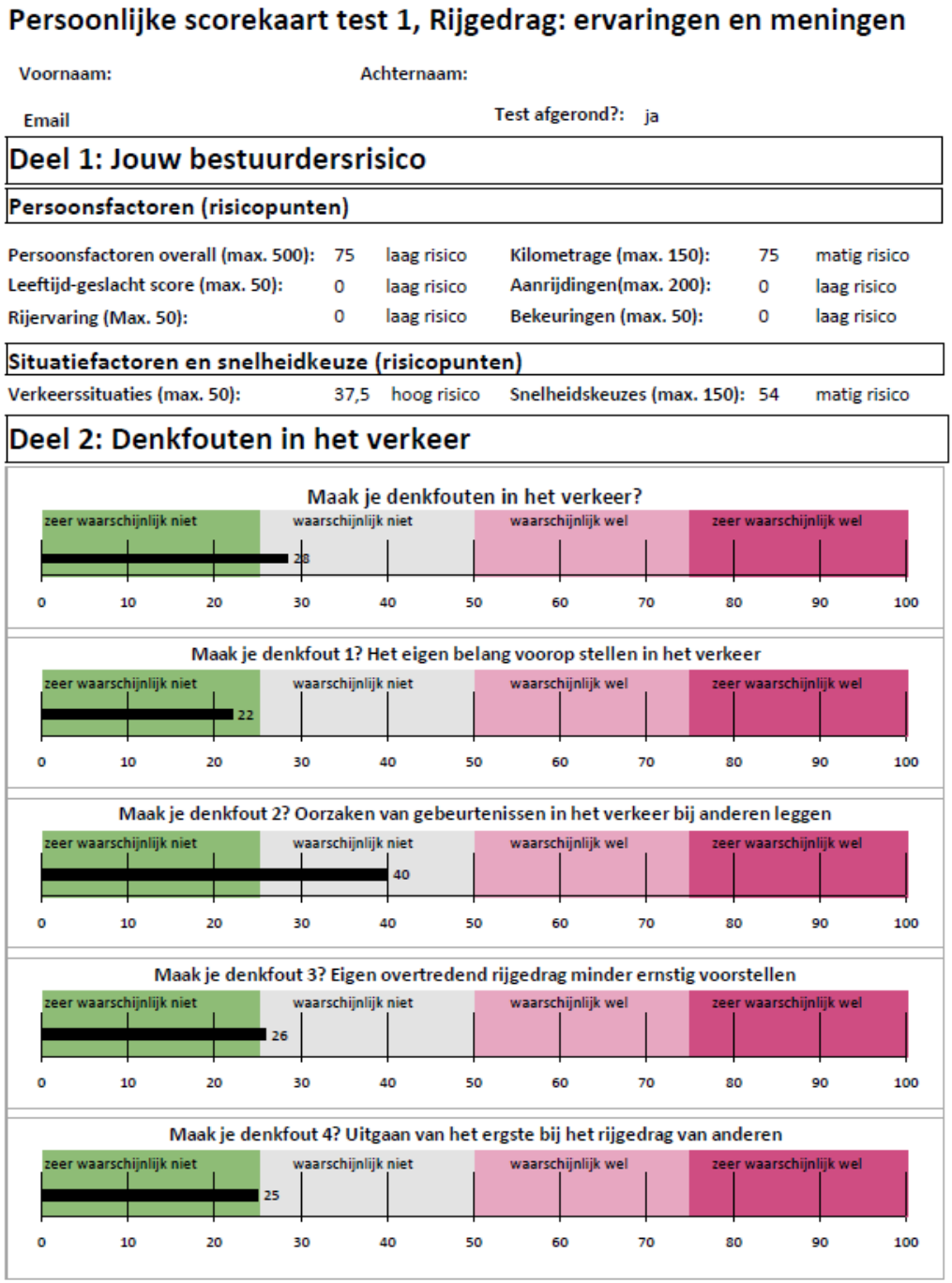
Scoring. The respondent has to choose an answer from a 6-point Likert-scale, ranging from "totally not agreed" to "totally agreed". No neutral answer can be given. Choosing "totally agreed" in most cases would result in a negative assessment result (high level of cognitive distortions). To prevent response bias a few 'positive fillers' (k = 5) are added. This means that the answer "totally agreed" would be positive instead of negative (low level of cognitive distortions) for the concerning items. These are included to prevent a responding pattern; gradually, respondents could conclude that "totally not agreed" would be the socially desirable answer, with the consequence that they would choose this answer without reading the item. Positive fillers would identify these respondents. An example of a positive filler: *Everyone should feel at ease in traffic.*

The items of the four categories and the positive fillers are in random order. Respondents could lose their attention if they got consecutive items of the same category. The variation in the sequence may prevent this. The sequence is shown in Table 5.

Table 5. Divisions of items per category

Category	Item
Self-Centered	1, 2, 8, 17, 19, 25, 27, 33, 38, 42
Blaming Others	3, 11, 12, 18, 26, 29, 34, 36, 39, 40
Minimizing / Mislabelling	4, 5, 9, 10, 23, 24, 31, 37, 41, 43
Assuming the Worst	6, 13, 15, 16, 20, 22, 30, 32
Positive fillers	7, 14, 21, 28, 35

Report. The respondents received their feedback reports after completing their assessments. An example of such a report can be seen in Figure 10.



Persoonlijke scorekaart test 1, Rijgedrag: ervaringen en meningen

Voornaam: Achternaam:

Email Test afgerond?: ja

Deel 1: Jouw bestuurdersrisico

Persoonsfactoren (risicopunten)

Persoonsfactoren overall (max. 500):	75	laag risico	Kilometrage (max. 150):	75	matig risico
Leeftijd-geslacht score (max. 50):	0	laag risico	Aanrijdingen(max. 200):	0	laag risico
Rijervaring (Max. 50):	0	laag risico	Bekeuringen (max. 50):	0	laag risico

Situatiefactoren en snelheidkeuze (risicopunten)

Verkeerssituaties (max. 50):	37,5	hoog risico	Snelheidskeuzes (max. 150):	54	matig risico

Deel 2: Denkfouten in het verkeer

Figure 10. Example of feedback report Assessment 1: traffic situations, opinions & experiences

The first part shows someone's driving risk, based on the answers given in the first part of the assessment, starting with the overall score, followed by the individual scores (age-gender, driving experience, driven miles, accidents and fines). For this report some assumptions were made, based on earlier research. Besides information coming from the conceptual framework, two principal researchers in the field of traffic psychology were asked about cut-off points for assigning risk points to different risk categories regarding the concerning variables. Based on the results of both methods, the framework of the risk factors was established. For example: men have more driving risk than women, as well as people with little driving experience and a high mileage. The factor 'Situations' is about dangerous situations, which can be influenced by specific times and locations; highways are safer than rural roads, and driving during weekend nights is more risky than during the week in the daytime.

The second part of the report is about the second part of the assessment, the self-serving cognitive distortions. The report indicates which type of cognitive distortion belongs to someone's behaviour.

Each item was scored on a 5-point scale. For "totally not agreed", zero points were given, whereas "totally agreed" resulted in five points. The intermediate values were subsequently worth one, two, three, and four points. After scoring the items, the total score of a scale was calculated, and then the sum was divided by the amount of items in the subscale (eight or ten). This outcome was multiplied by 100, which resulted in a score between 0 and 100. The indications of the scores are presented in Table 6:

Table 6. Meaning scores assessment 1: Cognitive distortions (Traffic situations)

Score	Meaning
Higher than 75	Most likely applicable
Between 50 and 75	Possibly applicable
Between 25 and 50	Probably not applicable
Lower than 25	Definitely not applicable

The first score shows whether there are cognitive distortions in traffic in general. The green area means that there are very likely no cognitive distortions in someone's driving behaviour. In the example feedback report there is a score of 28, which ends up in the light green area (*probably not*). This means there is a score of 28 out of 100. Since there is the possibility that someone has a certain type of cognitive distortion, the general score is followed by scores on the different types. In the example can be seen that the highest score (40) can be found on type 2: Blaming others.

Assessment 2: Self-evaluation of driving style. In Appendix 2 a print screen of Assessment 2 can be found. In this assessment about empathy with other road users it is studied whether someone in a few traffic situations thinks mostly about himself, or about others (Gibbs et al., 1992). In other words, if someone acts out of self-interest or in the interest of someone else, and to what extent he or she takes the other into account in the choices that have to be made. Important variables in this context are empathy, the (worry about the) well-being of others, but also having the knowledge of and complying with the traffic rules.

Eighteen common everyday traffic situations are presented. Every situation contains a possible space conflict for which a social solution is possible. This way, a moral consideration is asked. The respondents are shown a picture of a traffic situation. When taking into account the moral decision model (Figure 2), the next steps are included in this assessment: by means of this direct observation, the respondent could try to empathize with the situation (*perceiving/empathizing*). Based on the perception, the respondent should consider the different possibilities as solution (*reasoning/considering*). Every situation offers three possibilities, ranging from an altruistic choice to an egocentric one. These steps are not directly assessed, but are still necessary to take before the subsequent steps, which are actually assessed: the respondent needs to decide which solution he chooses (*judging/decide*) and how he will act (*acting*). Based on the decision and the act of the respondent, a certain effect will occur.

In Table 7 the division of the different types of space conflict of this assessment is shown. The most frequent type is 'others in an unfavourable situation' ($k = 11$).

Table 7. Types of space conflict in traffic situations

Type of space conflict	Frequency
Others in an unfavourable situation	11
Others taking your space	5
Common space	2

In Table 8 the division of different types of solutions for a space conflict is presented. The most frequent type of solution is 'allot space' (k = 21), followed by 'take space' (k = 14).

Table 8. Types of solution for space conflicts in traffic situations

Type of solution for a space conflict	Frequency
Allotting space	21
Taking space	14
Waiting and seeing how the other will do	11
Claiming space	10

The different traffic situations contain different driving tasks the driver has to perform in that moment. In Table 9 the division is presented of these tasks. The most frequent asked driving task is cruising (k = 10). Again, only the present tasks were included in the table. Examples of excluded tasks are: driving backwards, turn around, and stop.

Table 9. Types of driving tasks in traffic situations

Type of driving task	Frequency
Cruising with other traffic	10
Crossing a roundabout	2
Meeting oncoming road users (opposite direction)	2
Passing other road users	2
Crossing an intersection	1
Overtaking other road users	1

Also in this assessment the situations are based on common violations or annoyances in traffic (Infonu, 2014). In all situations no violation is needed for a social solution. An example item:

> *You look through the front window of your car. Right in front of you two cyclists are overtaking a parked truck. What do you do?*
>
> *a. I overtake the cyclists*
> *b. I hit the brakes and stay behind the cyclists*
> *c. I approach the cyclists*

Answer "a" indicates an egocentric perspective, whereas "b" indicates an altruistic one. Answer "c" is more neutral; no one benefits from this choice (perhaps the cyclists get annoyed or feel chased by the approaching car). A high level of thinking about others is related to the desire to place the interests of others above self-interest, but also the ability to empathize with others and being able to oversee the consequences of own choices for others. A high level of thinking about yourself may indicate a lack of empathy and placing self-interest above the interest of others. Acting out of self-interest could also have to do with just complying with the traffic rules, since it could lead to confusing situations for others by not complying with them. However, if the rules are not complied with or only in cases it is beneficial for yourself, it could indicate a negative attitude towards the rules.

Scoring. The answers are scored by means of the marble method, as described in the evidence model. This way, he or she does not have to agree totally with one option, and could give two options the same amount of points, for instance. In Figure 11 is the example item as included in the assessment presented. This example item was derived from the similar assessment for cyclists (which is not included in this study). It is emphasized that the total amount of divided points should count to 100.

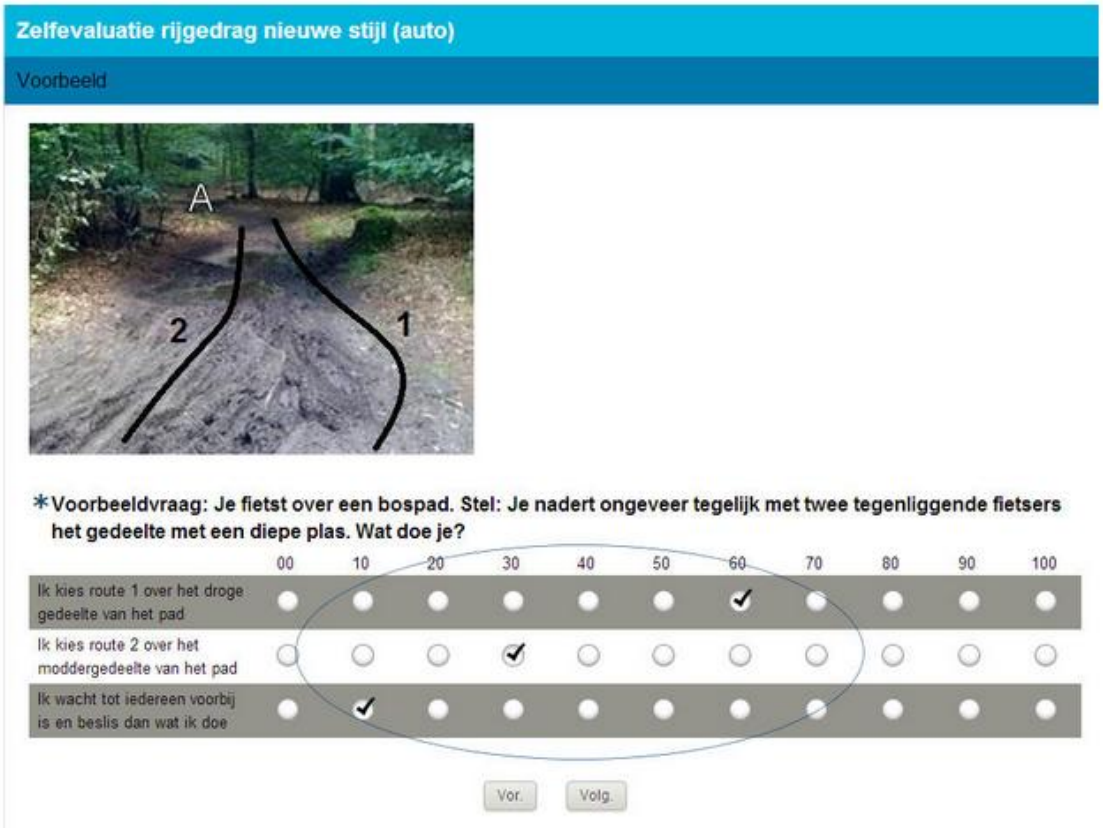


Figure 11. The example-item of Assessment 2: Self-evaluation of driving style

Report. Figure 12 shows the feedback report of the second assessment, the self-evaluation of someone's driving style.

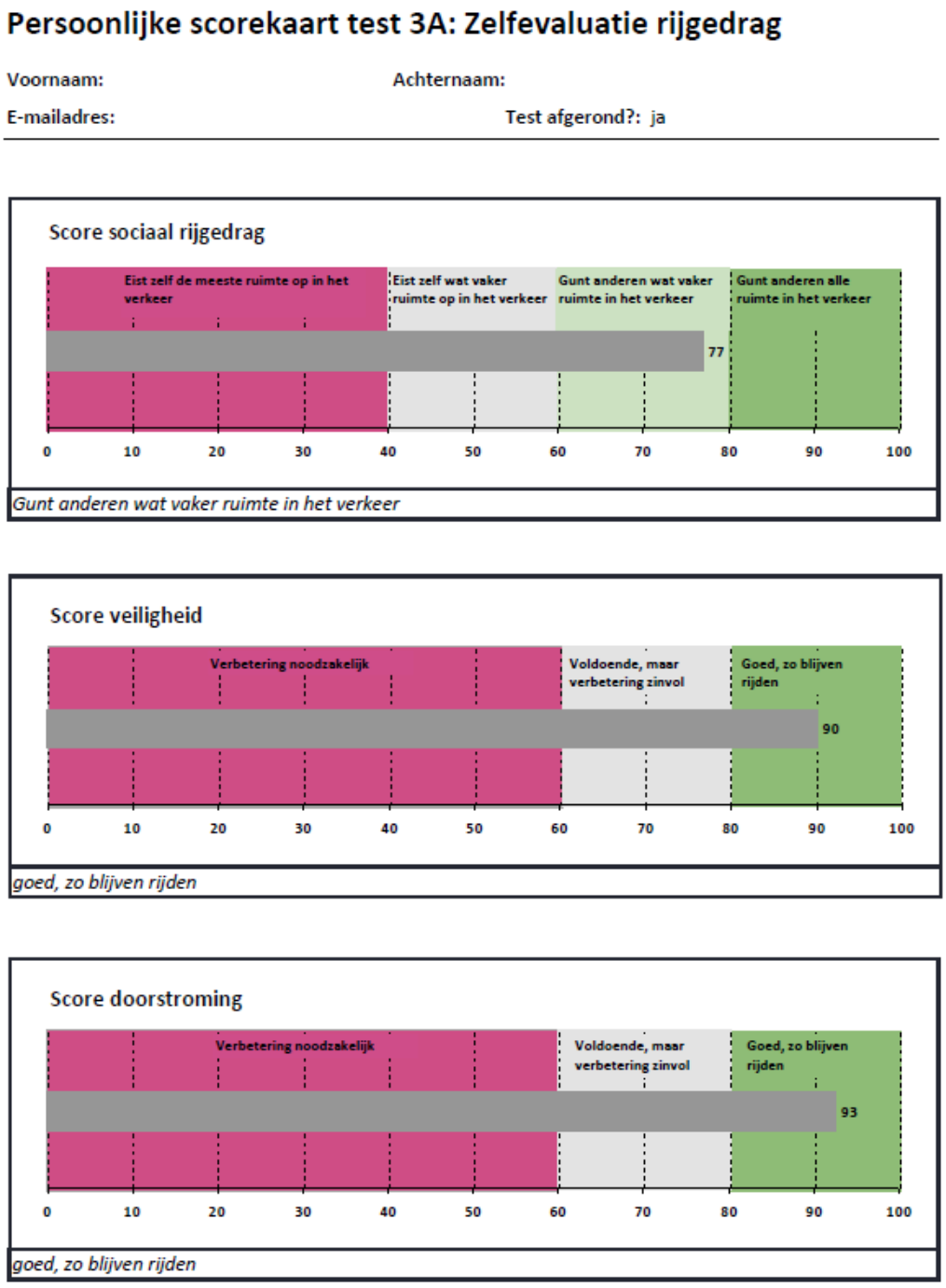


Figure 12. Example of feedback report of Assessment 2: self-evaluation of driving style

The first block in the report is about empathy. For every item, three solutions were possible. For the solution 'take space', zero points were given. For 'allot space', the respondent got two points. The in-between, more neutral option resulted in one point. This neutral option was only possible in eleven of the eighteen items, which means that seven items only provided an egocentric or altruistic option.

For all eighteen items the assigned percentage per option was summed up and weighted and divided by the maximum amount of points possible (36). The individual scores of the three options were added up to one score. This amount resulted in a score between zero and hundred.

For the marble method, 100 points had to be divided. Sometimes, however, this went wrong. When the total amount of assigned percentages for an item exceeded 100, the percentages per option were adjusted, assuming that the total amount should have been 100. For example:

Wrongly assigned percentages	
Option 1 (egocentric):	40
Option 2 (altruistic):	70
Option 3 (in-between):	10
Sum:	120%

Corrected percentages	
Option 1: 100/120*40 =	33.3
Option 2: 100/120*70 =	58.3
Option 3: 100/120*10 =	8.3
Sum:	100%

In Table 10 the meaning of the scores are presented.

Table 10. Meaning score assessment 2: Empathy (Self-evaluation)

Score	Meaning
Higher than 80	Generally allots space to other road users
Between 60 and 80	Allots space more often than taking space at the expense of others
Between 40 and 60	Takes space more often than allotting space to other road users
Lower than 40	Generally takes space at the expense of other road users

The second block of the report is about safe behaviour in traffic. Respondents had to indicate how often they commit a certain mistake that could lead to unsafe situations (e.g. overlooking other road users, reacting too late, being surprised). Furthermore items were included about taking safe routes (taking the ring road instead of driving through a residential), and about keeping distance and complying with the safety limits. Also here the scores vary from zero to hundred. The meaning of the scores are presented in Table 11.

Table 11. Meaning scores Assessment 2: Safety

Score	Meaning
Higher than 80	Improvement necessary
Between 60 and 80	Sufficient safe, but improvement is useful
Lower than 60	Good and safe

The example respondent scores 90 on this block. A high score means that someone is able to recognize dangerous situations (such as approaching traffic), that he is aware of everything that happens on the road and in the environment, that he notices changes in the traffic situation, that he takes care of safety margins to be able to act in time (such as following distance), that he adjusts his driving speed to the situation, and that he allots space to others to perform actions.

For the last block about traffic flow, respondents had to indicate again how often they commit a certain mistake that leads to hindering others or impeding the traffic flow of others (e.g. if others

have to hit the brakes when you merge, or when you block an intersection by not driving). The scores vary again from zero to hundred. The meanings are presented in Table 12.

Table 12. Meaning scores Assessment 2: Traffic flow

Score	Meaning
Higher than 80	Improvement necessary
Between 60 and 80	Sufficient fluent, but improvement is useful
Lower than 60	Good and fluent

The example respondent scores 83 on this block. A high score indicates that he is able to contribute positively to the traffic flow. He acts decisively in traffic, performs actions fluent, does not let other road users wait unnecessary, makes optimal use of the different driving lanes, does not hinder others if he has to stop or wait.

Assessment 3: Motives for driving style. In Appendix 3 print screens of Assessment 3 can be found. This assessment is about moral justification, about the awareness of the consequences some choices and actions have. It is about evaluating reasoning by means of moral criteria, to be able to determine if and why something is morally right or wrong. In this assessment the level of moral justification is assessed by means of the stages of Gibbs et al. (1992) in Table 2. These four stages can be divided in two driving styles, as can be seen in Table 13.

Table 13. Driving behaviour in combination with justification and description of Gibbs et al. (1992)

Justification to comply with the rules	Driving behaviour	Related level description of Gibbs et al. (1992): Focused on …
1. To prevent punishments or fines	Focused on yourself	… the physical (fine, punishment) and is short-lasting
2. To prevent unsafe situations for yourself		… own preferences and pragmatic needs, extrinsic motivated
3. To prevent unsafe situations for others	Focused on others	… the well-being of others, knowing the consequences of own reasoning
4. To prevent unsafe traffic or impeding traffic flows		… social importance. General interest is more important than self-interest

It is emphasized that this assessment is not about what you should do, but what you would do if a similar situation occurred. Also it is mentioned that there are sometimes reasons to decide to ignore the traffic rules. Those reasons are important to the assessment results, since they can be classified in different levels, as stated at the beginning of this section (Gibbs et al., 1992). Acting on lower levels (wanting to prevent a fine or difficult situation for yourself) asks more for a change in behaviour than when you act on higher levels (wanting to prevent difficult situations for others or prevent unsafe traffic in general). Furthermore, it is emphasized that all answers stay anonymous. Only within the study it is known which respondent gave which answers, but solely the respondent self will know his or her results.

Similar to the second assessment, the self-evaluation, the items represent different attributes of the traffic situation. In Table 14 is shown which type of violation are applicable. The most frequent violation in these items is about the (wrong) position on the road ($k = 3$).

Table 14. Type of traffic violation

Traffic violations	Frequency
(Wrong) position on the road	3
(Not) giving priority	1
Crossing railroads	1
Holding one hand on the wheel	1
Honking the horn	1
Ignoring traffic light	1
(Not) Indicating direction	1
Parking wrongly	1
Phoning while driving	1
Short following distance	1
Speed violation	1
Under influence of drugs while driving	1

Also, there were different driving tasks in these situations. These are shown in Table 15. Keeping track is the most frequent task ($k = 7$).

Table 15. Division of types of driving tasks

Type of driving task	Frequency
Cruising with other traffc	7
Crossing an intersection	2
Overtaking other road users	2
Exiting the road	1
Parking	1
Planning a route	1
Turning left at an intersection	1

In the first part of the assessment, respondents are presented seventeen times two situations with possible violations. The respondent has to choose which violation he thinks is more severe. The situations are partly based on most common annoyances in traffic (Infonu, 2014). In the second part, fifteen situations are presented. For each, a few questions have to be answered. Every situation contains a violation, either a legal traffic offense or a social offense. This way, a moral consideration is needed. The question is to what extent the respondent shows the presented behaviour and what justification he or she has for this behaviour.

According to the moral decision model (Figure 2), the next steps have to be taken: there is a direct perception in these situations, after which the respondent could try to empathize with the situation (*perceiving/empathizing*). This part will not be assessed. The parts that will be tested are the considerations the respondent has after his or her perception (*reasoning/ considering*), the decision that will be made after considering (*judging/deciding*) and, finally, the way he or she would act in the given situation (*acting*). Based on the decision and the way of acting, a certain effect will occur. An example item:

> *You approach (equivalent) crossroads. To your right, cyclists approach the crossroads too. Out of ten times, how often do you continue driving without giving priority?*

Then, two follow-up questions are presented:

- *In case you continue driving without giving priority, what are your reasons?*
  1. *The chance that I get caught is small*
  2. *It takes me less time*
  3. *Other drivers are not bothered by this*

- *In case you hit the brakes and do give priority, what are your reasons?*
  1. *I want to prevent getting caught or arrested*
  2. *I want to prevent ending up in a difficult situation*
  3. *I don't want to bother other drivers (discomfort, danger)*
  4. *I want to avoid causing traffic jams and unsafe traffic*

Naturally, when someone chooses "0" (zero) out of ten times at the first question, he will not get the first follow-up question. In theory, when the answer "10" (ten) is given, the respondent should only get the first follow-up question. However, the assumption has been made that this is not a realistic response, thus this question logic does not seem necessary.

Scoring. In the first part of this assessment the respondents have to choose which situation they think is more severe out of the two presented. An example of equation item can be seen in Figure 13. Situation 1 is about a long, abandoned road with a maximum allowed speed of 80 km/h. The violation contains driving 95 km/h or faster. In situation 2 the driver wants to from point A to point B. Route 1 is via the allowed traffic direction, while route 2 is against the traffic direction. The violation would be to choose route 2.

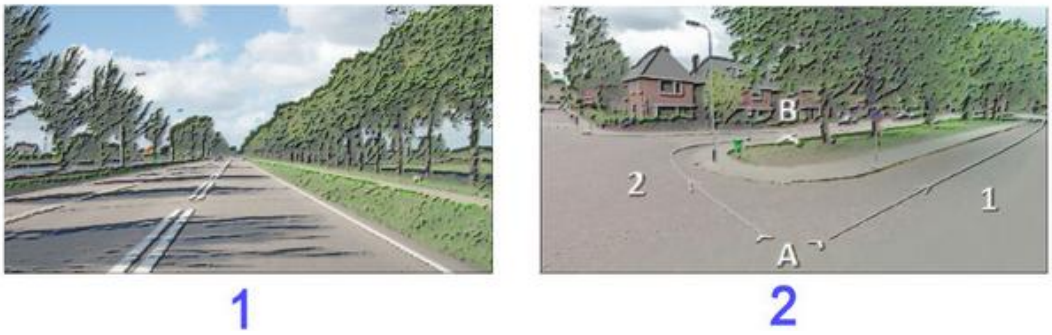


Figure 13. Example of an equation item of Assessment 3

Again, a 6-point Likert-scale is used, this time with the following options:

- 1 is much more severe
- 1 is more severe
- 1 is a bit more severe
- 2 is a bit more severe
- 2 is more severe
- 2 is much more severe

Beforehand it is established which of the two situations are actual more severe. An aspect that was used to decide about this severity, was the amount of a possible fine in the indicated situation. For the second part the 'marble' method was used again, where the respondent has to divide 100 points among the different answers.

Report. Figure 14 shows the feedback report of the third assessment: motives for driving behaviour.

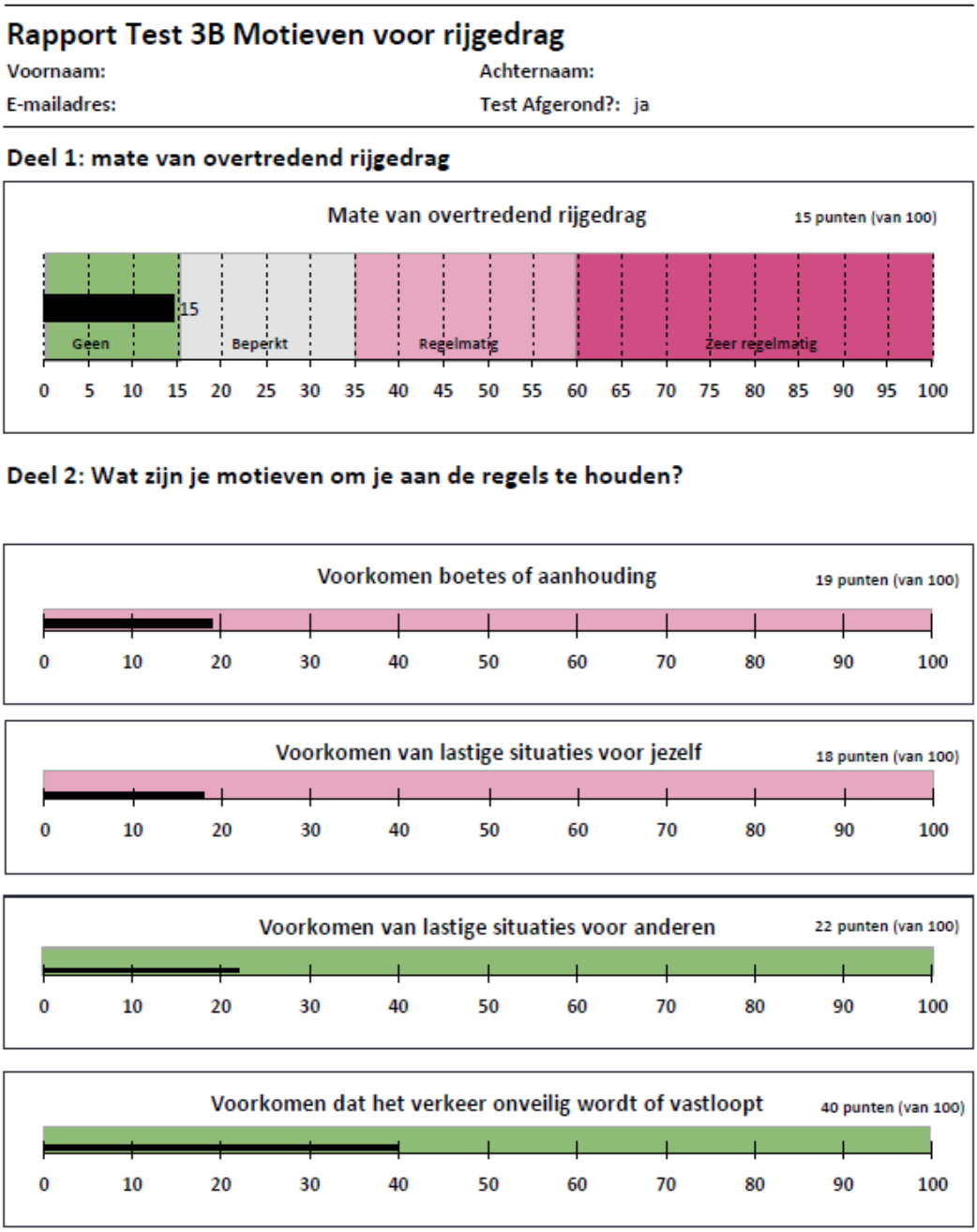


Figure 14. Example of feedback report assessment 3: Motives for driving behaviour

In this assessment is reported:

- The degree of violating driving behaviour
- The motives to comply with the rules

For the first part, the degree of violating and the gravity of the violation are combined to one score. In every situation the respondent can choose from zero to ten. The grave violation (k = 5) are worth three points per score. The fairly grave violations (k = 6) result in two points, and the not grave violations (k = 4) are only one point per score. This means that a maximum score of 150 + 120 + 40 = 310 can be scored. This is converted to a percentage score, with a maximum of 100. Table 16 shows the explanations for different scores.

Table 16. Explanations of scores on assessment 3: Justification

Score	Explanation
Lower than 15	No violating driving behaviour. In most cases you choose for omitting traffic violations and complying with the rules.
Between 15 and 35	Limited violating driving behaviour. Often you choose for omitting traffic violations. Sometimes you violate a rule.
Between 35 and 60	Regular violating driving behaviour. At times you commit a traffic violation. Sometimes it is a grave violation.
Higher than 60	Very regular violating driving behaviour. Often you choose for committing traffic violations, including grave violations as tailgating and driving on the emergency lane.

The respondent in the example feedback report got a score of 15, which is exactly on the margin of no and limited violating driving behaviour.

The second part is focused on the four levels of motives for complying with the rules. As already stated in Table 13, the first two levels of motives are aimed at the driver self, while level three and four are more aimed at other drivers. When scoring on the first two levels, it does not mean that someone's driving style is bad for traffic safety. However, if someone solely drives with these motives in mind, the driver does not see it as his job to promote the safety or comfort of other road users. If someone does comply with the rules because of wanting to prevent unsafe situations for others, he is strongly aimed at others and sees his driving behaviour as part of a larger system.

The scores on the four different levels sum up to 100. A higher score on one level indicates a lower score on another level. In the example feedback report can be seen that the respondent scores higher on level three and four, and thus lower on the first two levels.

Accompanying instruments

To study the predictive value of the new self-assessment for driving behaviour, the Driver Risk Assessment (Roelofs et al., 2013) are conducted as parts of the newly developed self-assessments. The selected questions relate to personal aspects as age, education, personal disorders and driver's licenses, and driving-related aspects as the chosen driven speed, the lane choice, the amount of fines, the amount of passive and active accidents, and the frequency of driving mistakes. The original DRA scored at sub-scales as 'Exposure to various risky traffic situations' (α = 0.78), 'Irritated towards others' (α = 0.63), 'Seriously angry towards other road users' (α = 0.51), 'Stays calm towards others' (α = 0.48), and 'Violation speed limits under various circumstances' (α = 0.92). Also, as mentioned before, the existing How I Think (HIT)-questionnaire was used (Brugman et al., 2011; Kuhlemeier et al., 2011), which is already discussed in detail at the corresponding assessment, Assessment 1.

Respondents

The respondents are derived from the DriveXperience, a program meant for young drivers between 18 and 25 years old, who hold their driver's license for one year. This resulted in a total of 755 respondents, which consisted of 396 men (52.5%) and 359 women (47.5%). The average age of the respondents is 20 years, as can be seen in Table 17. The high standard deviation might be explained by two respondents, who claim to have been born in 1927 and 1943. Another was born in 1981, while all other respondents were born between 1989 and 1997.

Table 17. Average age to gender

	Mean	n	SD
Male	19,7	396 (52.5%)	3.0
Female	20,0	359 (47.5%)	4.0
Total	19,8	755	3.5

Table 18 shows the educational level or work situation of the respondents in comparison with national data of CBS. It is remarkable that most respondents have a high educational level, which is considerably higher in the group of respondents than in the CBS-data (47.1% against 29.4%).

Table 18. Current education and work of respondents in comparison with CBS data

		Total amount	%
National (CBS)	Secondary education	220312	11.7%
	Vocational education (mbo)	407278	21.6%
	Higher education (hbo)	555146	29.4%
	Working or looking for work	706510	37.4%
Respondents (n = 754)	Secondary education	37	4.9%
	Vocational education (mbo)	230	30.6%
	Higher education (hbo)	354	47.1%
	Working or looking for work	130	17.3%

When looking at driver characteristics (Table 19), it can be seen how many kilometres the respondents have driven averagely, and the amount of crashes and fines was calculated. To make sure the statistics are in proportion, it was calculated how many crashes and fines the respondents would have had per million km.

Table 19. Mileage, crashes and fines

	Mean	SD
Mileage in 1000 km	6777.63	10193.71
Amount of active crashes per million km	50.98	215.87
Amount of passive crashes per million km	13.05	72.21
Amount of fines per million km	87.24	276.33

Data collection

The data collection took place during the fall of 2014, from the end of September till the end of December. All participants were invited by an e-mail of SurveyMonkey to complete the assessments. Also per e-mail they got their personal reports after completing the assessment, prior to their Coaching Day. This way the coaches could use the reports to plan their coached trips and adapt their approach to the individual needs, taking into account any striking features about the participants.

Data-analysis

In Table 20 the research questions and hypotheses are presented, in combination with the analysis that will be conducted to find answers on the questions.

Table 20. Research question and method of data-analysis

Research question	Analysis
1. Reliability assessment	Cronbach's alpha Correlations
2. Different levels of moral reasoning	Descriptive statistics
3. Relations between moral reasoning and other variables (*hypotheses*)	Correlations Regression analysis

The first question about the reliability of the three self-assessments will be studied by calculating Cronbach's alpha per assessment and per scale within the individual assessments. Also correlations between scales will be given. The second question about the levels of moral reasoning will be studied

by analysing descriptive statistics. To study the possible relations for the third research question, a (multiple) regression analysis will be conducted.

Results

Reliability and validity of instruments

The first research question addressed the reliability of the three developed assessments about moral reasoning. In Table 21 are the alpha's shown of the overall assessments and its subscales.

Table 21. Reliability of assessments

Assessment 1: Cognitive Distortions - Traffic situations	**Cronbach's alpha**
Overall assessment	.91
Scale 1: Self-centered	.70
Scale 2: Blaming others	.80
Scale 3: Minimizing / Mislabelling	.73
Scale 4: Assuming the worst	.67
Assessment 2: Empathy - Self-evaluation of driving style	**Cronbach's alpha**
Overall assessment	.72
Scale 1: Altruistic	.66
Scale 2: Egocentric	.72
Scale 3: In-between	.53
Assessment 3: Justification - Motives for driving style	**Cronbach's alpha**
Overall assessment	.82
Stage 1: Prevent punishment or fines	.86
Stage 2: Prevent unsafe situations for yourself	.79
Stage 3: Prevent unsafe situations for others	.76
Stage 4: Prevent unsafe traffic	.89

The most reliable (overall) assessment with a reliability of .91 is the first assessment, about cognitive distortions. The respondents' choices can be categorized in four subscales, which all have an acceptable reliability around an alpha of .70. The other overall reliabilities are also acceptable with alphas of respectively .72 and .82. The assessment about justification has the most reliable scales (.86, .79, .76, .89). The scale with the lowest alpha of .53 can be explained since it is a 'in-between' scale in the second assessment, which tends to measure the egoistic and altruistic aspects of the respondent. Since this in-between scale does not measure one of those aspects, a low reliability on this less concrete scale was therefore expected.

In Table 22 the correlations between the different assessment scales are presented. All categories of the cognitive distortions are positively correlated to each other, which was expected since they all address the lower moral stages. Furthermore, the altruistic perspective of the self-evaluation is negatively correlated to all cognitive distortions as expected, since the distortions are on a low moral lever, whereas the altruistic solutions indicate a higher moral level. In addition, the egocentric solution is positively correlated to the cognitive distortions, and thus negatively with the altruistic solution, since these are contrary. The first stage of Gibbs' moral model (preventing punishment) has a positive relation with the cognitive distortions and the egocentric solution, and a negative one with the altruistic solution. This stage is negatively correlated to the other three moral stages, so there seems to be a division between the first moral stage and the other three. These three 'higher' stages have a negative correlation with the cognitive distortions. This is remarkable, since it was suggested that the division would be between the lower two moral stages and the higher ones. However, stage 2 (wanting to prevent unsafe situations for self) barely has a correlation with all other variables; only with the first and fourth stage there is a clear negative correlation. Stage 3 (wanting to prevent unsafe situations for others) and stage 4 (wanting to prevent unsafe traffic) have no correlation at all.

In short, the correlations give an expected indication of the variables of moral reasoning. It is clear which variables indicate a low moral level, and which ones point to a higher one.

Table 22. Correlations between assessment scales

	1. Self-centered	2. Blaming others	3. Minimizing	4. Assuming the worst	All cognitive distortions	Allotting space	Taking space	1. Prevent punishment	2. Prevent unsafe situations for self	3. Prevent unsafe situations for others	4. Prevent unsafe traffic
1. Self-centered											
2. Blaming others	.64										
3. Minimizing	.73	.65									
4. Assuming the worst	.60	.67	.64								
All cognitive distortions	.86	.88	.87	.83							
Allotting space	-.41	-.32	-.38	-.28	-.40						
Taking space	.42	.35	.45	.35	.45	-.68					
1. Prevent punishment	.34	.31	.35	.31	.38	-.16	.23				
2. Prevent unsafe situations for self	-.04	-.08	-.10	-.06	-.08	-.09	.03	-.31			
3. Prevent unsafe situations for others	-.19	-.20	-.21	-.20	-.23	.15	-.12	-.58	-.05		
4. Prevent unsafe traffic	-.25	-.18	-.20	-.19	-.24	.14	-.21	-.56	-.42	.04	

Levels of moral reasoning

For the second research question is studied which levels of moral reasoning are perceived among young drivers of the DriveXperience. This question was divided in sub-questions.

First, it is studied to which degree young drivers justify norm-complying behaviour according to the different levels of moral reasoning. In Table 23 are the average scores per stage of moral reasoning presented. To justify norm-complying behaviour, in most cases the respondents would reason according to the second stage: to prevent unsafe situations for themselves, with an average of 28.7. This manner of justification is followed by the third stage: wanting to prevent unsafe situations for others (27.3). The fourth category follows then, with an average of 25.7. The remaining category has a much lower average score than the other three, which is 18.3.

Table 23. Moral reasoning levels and averages

			Reach	
Subscale	**Mean**	**SD**	**Min**	**Max**
1. To prevent punishments or fines	18.2	14.7	0	87
2. To prevent unsafe situations for yourself	28.7	9.9	0	66
3. To prevent unsafe situations for others	27.3	8.5	4	55
4. To prevent unsafe traffic or impeding traffic flows	25.7	12.0	4	84

Secondly, it is studied to which degree young drivers justify antisocial behaviour, and to what extent they use cognitive distortions for this. It is also studied if specific types of cognitive distortions are used. In Table 24 can be seen which average belongs to the different types of cognitive distortions. The respondents had to score their answer to a 6-point Likert-scale. These scores were converted to a scale of 0 to 100. The average score of the overall assessment is 26.7.

Table 24. Cognitive distortions and averages

			Reach	
Subscale	**Mean**	**SD**	**Min**	**Max**
All subscales	26.7	11.3	0.0	74.2
1. Self-centered	24.2	12.2	0.0	73.3
2. Blaming others	33.5	15.1	0.0	88.0
3. Minimizing / Mislabelling	23.6	13.2	0.0	91.1
4. Assuming the worst	29.5	13.7	0.0	72.5

The respondents disagreed the most with the cognitive distortion items of the category ‘minimizing’, with an average score of 23.6. This scale was followed by the self-centered category, with an average score of 24.2. Next category is ‘assuming the worst’, with a score of 29.5. The category that scored the highest, is ‘blaming others’, with 33.5. The respondents disagreed less with the items of this subscale.

Thirdly, it was studied to which degree young drivers allot space to other drivers, and to what extent they take space at the expense of others. In Table 25 and Table 26 is presented how many respondents chose for the egocentric (taking space), altruistic (allotting space) or in-between option. It can be seen that an average of 53.5% would choose for an altruistic solution in the given traffic situations (TS). That is more than three times as much as the average of respondents that would choose an egocentric solution (16.5%). The other 30.0% would choose an option that cannot be classified as egocentric or altruistic. In these solutions the respondent would wait and see what happens before he decides what to do.

Table 25. Allotting and taking space

Traffic Situations	Allotting space	Taking space	Wait and see
Car wants to merge with your lane	86.8	13.2	-
Approaching cyclists and child with ball	89.2	10.8	-
Slower cars want to merge with highway	85.3	3.6	11.1
Two cyclists overtaking parked truck	80.7	19.3	-
Truck suddenly goes left to your lane	81.6	18.4	-
Mother and child on bicycles in front of you	80.7	19.3	-
Car approaches from left	76.3	23.7	-
Road narrowing at same time	73.5	26.5	-
Road narrowing behind other car	50.9	10.8	38.3
Car stuck behind truck on highway	47.4	12.2	40.4
Cyclist wants to cross the crosswalk	45.4	8.2	46.4
Mother with child and a van	40.6	8.8	50.6
Being overtaken on narrowing street	28.5	8.9	62.6
Cyclist waiting to cross the road	29.2	8.0	62.8
Car drives from parking bay	19.9	17.2	62.9
Talking person with dog in the way	18.5	11.2	70.3
Other car in parking bay	17.1	44.5	38.4
Car wants to go on roundabout	11.1	33.3	55.6
Average	53.5	16.5	30.0

Table 26. Egocentric and altruistic behaviour

			Reach	
Subscale	**Mean**	**SD**	**Min**	**Max**
Altruistic behaviour (allotting space)	53.5	10.1	22.8	90.6
Egocentric behaviour (taking space)	16.5	9.6	0.0	56.1

In the situations where the 'wait and see'-option was not chosen, it was also not possible to choose a solution that could be classified as such. In six situations (7, 8, 13, 14, 16 and 18), the 'wait and see'-option was chosen by more than half of the respondents. In these situations being moral sometimes meant that the respondent could not comply with the rules.

Relations between moral reasoning and driving behaviour

For the third research question it was studied to what extent moral reasoning is related to driving behaviour, such as speed choice, complying with the rules, traffic violations, and accident involvement. Furthermore, it was studied to what extent there is a relation with the personal characteristics of gender and age.

In this case, seven separate multiple regression analyses were executed, whereby the direct effects of the predictors were established. The predictors are assumed to causally precede the concerning depending variable. The analyses started with a regression analysis with as depending variable the amount of involvements in accidents. By using the stepwise method the other blocks of predictors were added:

- Background variables: gender, age, mileage, moped experience
- Traffic experience: exposure to traffic
- Cognitive distortions
- Motives to comply with the rules
- Space-taking behaviour
- Driving speed on diverse roads
- The degree of violating behaviour

As a next step, the direct predicting variable followed – which precedes according to the conceptual framework, as included as depending variable in a regression analysis: degree of violating behaviour, as measured in the assessment about motives for driving style. This way the consecutive regression analyses were continued, with as depending variables driving speed, space-taking behaviour, two norm-complying motives which represent opposite moral levels (preventing punishment and fines, versus preventing unsafe traffic) and cognitive distortions (overall score). In Table 27 to Table 33 the related regression tables are presented.

Table 27. Test score for cognitive distortions predicted by background variables

	Beta	t	Sig
Constante		21.43	.000
Gender (coded women=0; men=1)	.14	3.44	.001
Years of experience on moped	.13	3.23	.001
Motive: prevention of fine´	.34	8.21	.000

Adjusted R-square= .18

Table 28. Test score for motive: prevention of fines predicted by background variables

	Beta	t	Sig
(Constant)		5.03	.000
Gender (coded women=0; men=1)	.14	4.10	.000
Exposure to risky traffic situations	.13	3.63	.000

Adjusted R-square= .07

Table 29. Test score for motive: prevention of fines predicted by background variable

	Beta	t	Sig
(Constant)		41.81	.000
Mileage	-.11	-2.46	.014

Adjusted R-square= .01

Table 30. Driving speed predicted by background variables and moral reasoning scales

	Beta	t	Sig
Constant		-4.66	.000
Gender (coded women=0; men=1)	.16	4.43	.000
Years of experience on moped	.10	2.78	.006
Test score for cognitive distortions	.43	10.39	.000
Space taking behavior at the expense of others	.12	3.04	.002

Adjusted R-square= .32

Table 31. Space taking behaviour at the expense of others predicted by background variables and moral reasoning scales

	Beta	t	Sig
(Constant)		5.77	.000
Test score for cognitive distortions	.45	8.71	.000
Exposure to risky traffic situations	-.11	-2.24	.026
Test score for motive for rule compliance: prevent that the traffic system fails	-.11	-2.28	.023

Adjusted R-square= .22

Table 32. degree of violating driving behaviour predicted by by background variables. moral reasoning scales. space taking behaviour and driving speed

	Beta	t	Sig
Constant		1.28	.200
Gender (coded women=0; men=1)	.14	3.22	.001
Test score for cognitive distortions	.19	3.37	.001
Test score for Motive: preventing fines	.18	4.03	.000
Space taking behavior at the expense of others	.11	2.35	.020
Driving speed	.33	6.38	.000

Adjusted R-square= .43

Table 33. Traffic fines predicted by variables preceding in the model

	Beta	t	Sig
Constant		1.44	.150
Mileage	.14	3.24	.001
Education level	-.11	-2.64	.008
Degree of violating driving behaviour	.27	6.43	.000

Adjusted R-square= .13

In Figure 15 the path-model of the conceptual framework is presented, with the results after performing the described consecutive recursive regression analyses. The green blocks are variables that relate to aspects of moral reasoning in traffic and for which results are provided. It can be seen that 44% of the variance in violating behaviour is explained by the set of variables which include moral reasoning. There are two direct effects on violating behaviour: the cognitive distortions for justifying antisocial behaviour, with a beta of .19, and the low moral level of justifying norm-complying motives (wanting to prevent punishment, such as fines), with a beta of .18. Some of the effects go via space-taking behaviour, or driving speed. The relation between space-taking behaviour and degree of violating is .11. The beta's between cognitive distortions and space-taking behaviour and driving speed are the highest in this model (respectively .45 and .43). There is also a relatively strong relation between driving speed and the degree of violating behaviour (.33). Between the motives of preventing fines and cognitive distortion also exists a relation of .33. The variable 'motives for unsafe traffic' has a negative relation of -.11 with space-taking behaviour, and space-taking behaviour has a relation of .12 with driving speed.

The beta's of the personal characteristics and the driving behaviour variables are less strong. The variable 'gender' has a beta of .14 with the degree of violating behaviour, a beta of .16 with driving speed, a beta of .14 with the cognitive distortions, and a beta of .18 with the motives for preventing fines. Age only has a (negative) relation with the motives of preventing fines (-.09), and exposure to traffic with cognitive distortions (-.11) and motives for preventing fines (.16). Moped experience has a negative relation with space-taking behaviour (-.11), and positive ones with driving speed (.10), and cognitive distortions (.13). At last, mileage has relations with involvement in accidents (.14), driving speed (.09) and a negative relation with motives for preventing unsafe traffic (-.11).

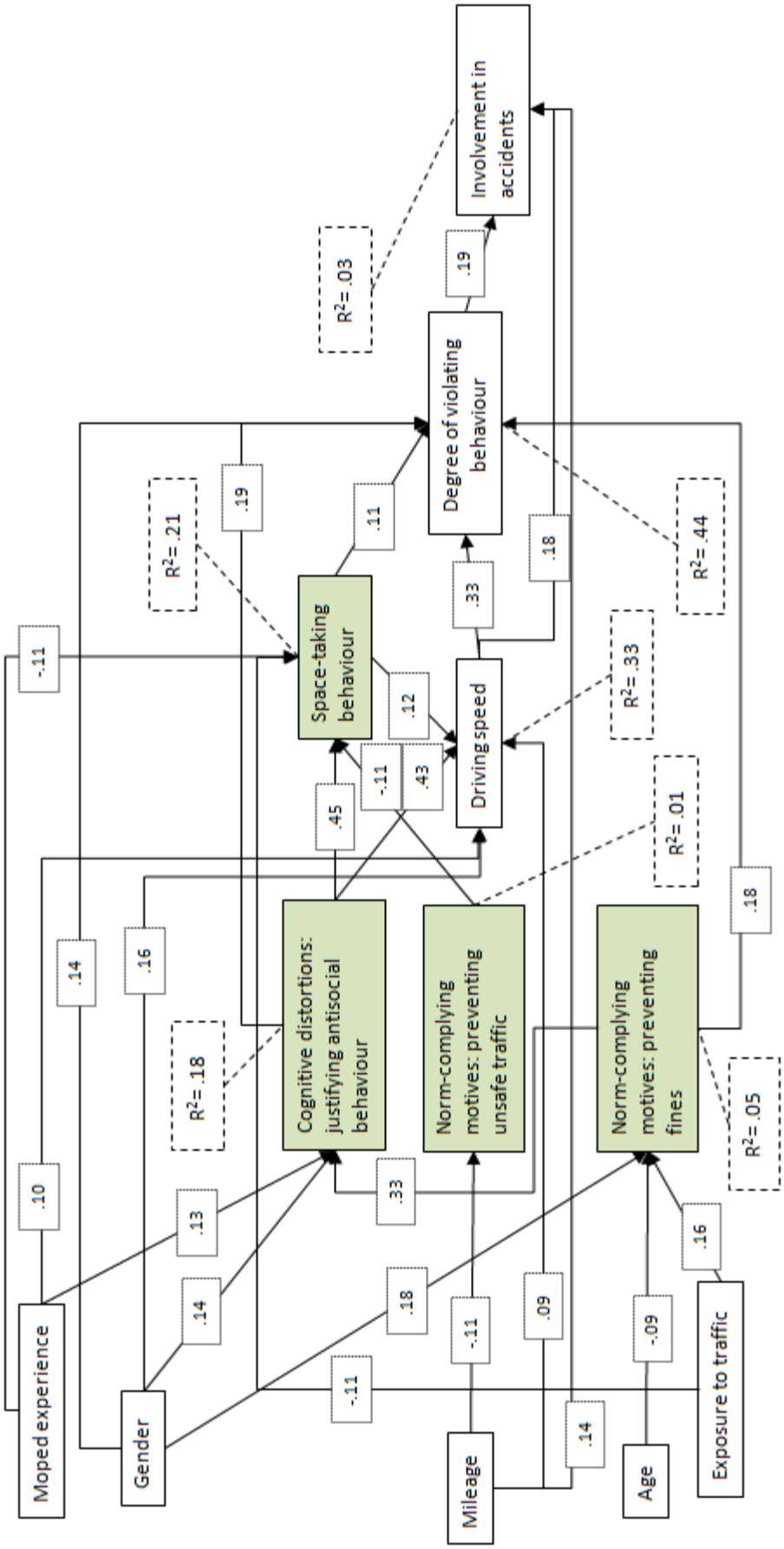


Figure 15. Path-analysis

Conclusion & Discussion

Summary of research questions and conclusions

Reliability and validity of developed assessments. The assessment's alpha's indicate that the reliability of the assessments are sufficient enough to work with. The second assessment – with the lowest overall alpha– contained a scale that could not be clearly defined, in contrast to the other two scales of egocentric and altruistic options. When looking at the alphas in **Fout! Verwijzingsbron niet gevonden.**, it can also be seen that this scale has a lower alpha than the other two. This scale is possibly the reason for the lower alpha. However, for the first research question can be concluded that the aspects of moral reasoning are established as reliable.

Levels of moral reasoning among the respondents. It was studied to which degree young drivers justify norm-complying behaviour according to the different levels of moral reasoning. The lowest moral level (preventing punishment or fines) was selected the least often of the four. The scores of the other three moral levels are closer to each other. This difference seems remarkable. It means that drivers do not in particular justify their norm-complying behaviour because they want to prevent punishment for themselves. Their motives are more directed to the safety of traffic, mostly about their own safe driving, but also about the safety of others and the safety of the whole traffic system. For this part can be concluded, based on these results, that the respondents justify their norm-complying behaviour on a relatively high moral level.

Furthermore, it was studied to which degree young drivers justify antisocial behaviour, and to what extent they use (specific types of) cognitive distortions for this. The average score on the assessment of cognitive distortions was 26.7 out of 100. This means that, in most cases, the respondents stated that they (totally) did not agree with the given statements. Therefore, it could be concluded that cognitive distortions are used to a small extent for justifying antisocial behaviour. To the least extent the category of minimizing and mislabelling was used. The most chosen category was blaming others. This means that, in case people use cognitive distortions, they would mainly blame others for their own antisocial behaviour.

For the last part of this research question it was studied to which degree young drivers allot space to other drivers, and to what extent they take space at the expense of others. Altruistic behaviour resulted in a score of 53.5%, while egocentric behaviour only got a score of 16.5%. Therefore, it can be concluded that the respondents were mostly altruistic, which means allotting space to others happens more than taking space at the expense of others. The remaining percentage did not show clear choices in behaviour (they would rather 'wait and see'). In the situations where this option was included, being moral sometimes meant that the respondent could not comply with the rules. For example, there was a situation where a cyclist was waiting at the middle of the road to cross. As a driver you could choose to let the cyclist go, instead of driving on. However, this could also mean that there is a change that you impede the traffic flow. Considerations like this could result in a 'wait and see'-decision.

Relations between moral reasoning and driving behaviour. The results showed that there are two direct effects on violating behaviour: cognitive distortions and the lowest level of moral reasoning (preventing punishment). Violating behaviour has a direct effect on the involvement in accidents, which indicates that lower levels of moral reasoning have an indirect effect on a higher number of accidents, as expected (*hypotheses a*). Most relations between moral reasoning and violating behaviour go via space-taking behaviour and driving speed. Space-taking behaviour is positively related to a higher driving speed. This seems logical, since both variables indicate a low moral level or some kind of violation. Also, it is confirmed that cognitive distortions are related to a higher driving speed (*hypotheses b*) and to a higher degree of space-taking behaviour (*hypotheses c*), which was also expected, since cognitive distortions only occur on the lower moral levels. The model confirms this with the fact that the lower moral motives of preventing fines are indeed related to a higher degree of cognitive distortions. In addition, it is indicated that higher moral motives (wanting to prevent unsafe traffic) are related to a lower degree of space-taking behaviour. This was also expected, since a low degree of space-taking behaviour indicates a more altruistic perspective.

Furthermore, the relation between the moral variables and personal characteristics was studied. It turns out that the male gender is related to a higher degree of violating behaviour (*hypotheses d*), a higher driving speed, and that men use cognitive distortions to a higher extent than women. Moreover, they justify their behaviour on lower moral levels than women. Besides, the results show that the younger the driver is, the higher the degree of justifying on lower levels will be. Indirectly (via low moral levels), age has a relation with a higher degree of violating behaviour (*hypotheses e*). The conclusions turned out as expected. In the assessments also the variables 'mileage', 'moped experience' and 'exposure to traffic' were included. It can be stated that the higher the mileage, the more the behaviour will be justified on lower moral levels, such as cognitive distortions and wanting to prevent punishment and unsafe traffic for yourself. A higher mileage also indicated a higher driving speed and a higher degree of involvement in accidents. For exposure to traffic it can be stated that the higher the exposure, again the more the behaviour will be justified on lower moral levels. However, in contrast: a higher exposure to traffic is related to a lower degree of space-taking behaviour. This is a remarkable result, since the variables on the lower moral levels seem to be connected in all other cases. It is possible that the drivers with a high exposure to traffic use the 'wait and see'-perspective more than an egocentric one.

It can be concluded that all hypotheses are confirmed. Driving speed, space-taking behaviour and violating behaviour are predictors for moral reasoning. For personal characteristics can be stated that men mostly show more risky driving behaviour then women. Also a higher mileage and a higher exposure to traffic are predictors of lower levels of moral reasoning. Furthermore, the younger the driver, the bigger the chance of a riskier driving style than that of older drivers. These outcomes were also stated in the literature (Deery, 1999; Ryan, Legge & Rosman, 1999; Al-Balbissi, 2003; Yagil, 1999).

Discussion

For the present study it was decided to conduct three different assessments about moral reasoning, to try to measure the respondents' levels of moral reasoning for different variables. Completing all three assessments require a lot of time of the respondent. However, shorten the assessments could have implications for the reliability and validity of the assessments. Though, since two assessments (about justification and empathy) are scaled on the lower and higher moral levels, and one assessment (cognitive distortions) are scaled on just the lower levels, it might be considered to establish the respondents' level of moral reasoning by means of the first two assessments. The expectation is that only a small percentage of the respondents would score in the critical area of moral reasoning, an expectation based on results of other studies (Barriga et al., 2001; Kuhlemeier et al., 2011). Only in case a respondent shows low moral levels above average, it might be useful to let him or her complete the assessment about cognitive distortions. This saves time and possible boredom of respondents with higher moral levels.

In addition to the mentioned above, it should be taken into account that the established scores are dependent of the accuracy of the measurement, and not so much of the measurement itself. Furthermore, the reported scores are based on the respondent's perceptions of his or her own driving behaviour, which means it could not be proven if the data agree with the reality. Therefore a validation to higher standards are necessary.

For the respondents it was possible to leave a comment at the end of every assessment. This was meant as evaluation, but some respondents used it to complain about individual items or the overall assessments. These comments could come in handy to consider the assessments. For example: a considerable amount of respondents thought the assessment took too much time. They were not prepared for that; it might be useful to mention the needed time in the invitation mail, or the assessments could be adjusted a bit (less answer possibilities, less items). Or, as stated in the first discussion point, not all respondents should have to complete all three assessments.

For the scoring method the 'marble method' was used. With this technique it was possible to let the respondents consider the different answer possibilities in detail, and it was not necessary to totally agree with one of these options. This method was chosen to yield a representative picture, as much as possible. Naturally, it is possible that another method would be more accurate or reliable. However, another method was not considered for the present study.

It turned out that most variables had a (strong) relation with each other. However, this is not surprising. Most variables were intertwined in the assessments, which is a logical cause for the high correlations. Therefore it would have been more representative if the different variables were only asked in individual assessments, instead of in every assessment.

The assessment were conducted on the online platform of SurveyMonkey, which had its limitations. For the marble method it would have been useful if respondents could assign no more than 100 points to the different options. However, they had to count for themselves how many points they had divided yet. A counter would help in this case, since the respondent would then be able to see how many points were left to divide. SurveyMonkey had to be chosen since it is a free tool; there are probably possibilities with other online tools.

The assessment for cognitive distortions maintains a 6-point Likert-scale, in accordance with Gibbs (1991). Before that was decided, it was considered to choose 4, 6 or 10 possibilities. A larger scale could give more information, but for the respondents it gets harder to distinguish the different possibilities then. This also depends on age: for youth a 5- or 6-point Likert-scale is recommended (Morrison, Ross & Kemp, 2007). Therefore, a 6-point Likert-scale was chosen eventually.

In this study also the variable 'behavioural disorder' was included. Eventually it did not seem to fit between the other categories of variables. Behavioural problems may have some influence on the interaction in traffic and in the car with other passengers, but some types are not expected to have any influence on the results of the assessments. Therefore it turned out to be a complex variable. Eventually it was chosen to not include this variable in the analyses.

Also, age as variable was studied. However, the target group only contained young drivers between 18 and 24 years old. This made the variable limited. Therefore, the conclusions about age (younger drivers show in general more risky behaviour than older drivers) only can be applied for this age group of young drivers.

Practical consequences

This study confirmed that moral reasoning and empathy are aspects that influence someone's driving style. Therefore it could have beneficial effects if driving schools would try to emphasize these aspects more in their training, by focusing more on socio-moral self-reflection. Important in this context is to ask their young drivers questions such as what their role in traffic is and how their interests relate to those of others and the whole traffic system. In the end, this might contribute to lower accident rates in the target group of young drivers.

Furthermore, performing in this context turns out to be not only a matter of someone's ability and knowledge, but also a matter of their attitude and wanting to perform in a certain way. This should be emphasized in (driver) training and education.

Perspective for future research

In this study a specific target group was chosen: young people between 18 and 24 years old who just held their driver's license, who participated in the DriveXperience. It could be interesting to let other target groups complete the assessments, to find out if there are some remarkable differences between different target groups. Elderly people might be more careful in traffic, while lease drivers may care less about the safety of their car. Furthermore, it turned out that the largest part of the respondents in this study was highly educated. A more representative sample deliberately involving more lower educated people is desirable.

Future research to this topic could be in the form of a combined study. The quantitative part of this study together with a qualitative part could result in more reliable and valid conclusions. The qualitative part could consist of interviews with a sample of the respondents to confirm some findings. By interviewing, it is possible to go deeper in causes, motives and processes when talking with an individual.

Reference list

Al-Balbissi, A.H. (2003). Role of Gender in Road Accidents. *Traffic Injury Prevention, 4* (1), 64-73

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association

Baarda, D.B., Goede, M.P.M., de, & Teunissen, J. (2009). *Basisboek kwalitatief onderzoek*. Groningen/Houten: Noordhoff Uitgevers.

Barriga, A.Q., Gibbs, J.C., Potter, G.B. & Liau, A.K. (2001). *How I Think (HIT) Questionnaire Manual.* Champaign, II.: Research Press

Beerthuizen, M.G.C.J. (2012). *The impact of morality on externalizing behavior: values, reasoning, cognitive distortions and identity*. Utrecht: Zuiddam Uithof Drukkerijn

Berk, L.E. (2006). *Child development* (7th edition). Boston: Pearson

Bianchi, A. & Summala, H. (2002). Moral Judgment and Drivers' Behavior among Brazilian Students. *Psychological Reports, 91,* 759-766

Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kapan, 80* (2), 139-148

Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 1* (1), 1-40

Blasi, A. (1980). Bridging moral cognition and morel action: A critical review of the literature. *Psychological Bulletin, 88,* 1-45.

Brugman, D., Nas, C.N., Velden, F. van der, Barriga, A.Q., Gibbs, J.C., Granville, B.P. & Liau, A.K. (2011). *Hoe Ik Denk Vragenlijst (HID) Handleiding.* Amsterdam: Boom test uitgevers

Christie, R. (2001). *The effectiveness of driver training as a road safety measure: An international review of the literature.* Road Safety Research, Policing and Education Conference, Melbourne, Victoria, Australia

Colby, A. & Damon, W. (1992). *Some do care: Contemporary lives of moral commitment.* New York: Free Press

Craen, S. de (2010). *The X-factor: a longitudinal study of calibration in young novice drivers* (Doctoral dissertation, Technische Universiteit Delft). Retrieved from http://repository.tudelft.nl/view/ir/uuid:2083b0dc-e597-4b8b-908a-b49ed9f614ae/

Dahlen, E.R., Martin, R.C., Ragan, K. & Kuhlman, M. (2005). Driving anger, sensation seeking, impulsiveness, and boredom proneness in the prediction of unsafe driving. *Accident Analysis and Prevention, 37,* 341-348

Dam, G. ten & Volman, M. (2000). Sociale competentie: reddingsvest en levenskunst. Over de pedagogische opdracht in de praktijk. *Pedagogiek, 20* (2), 112-127

Deery, H.A. (1999). Hazard and Risk Perception among Young Novice Drivers. *Journal of Safety Research, 30* (4), 225-236

Elander, West & French (1993). Behavioral correlates of individual differences in road-traffic crash risk: An examination of methods and findings. *Psychological Bulletin, 113* (2), 279-294

Emmerik, M. L. van (2004). *Beyond the Simulator. Instruction for High-Performance Tasks.* Soesterberg: TNO.

Fuller, R. (2005). Towards a general theory of driver behaviour, *Accident Analysis and Prevention, 37,* 461-472

Garrison, C., & Ehringhaus, M. (2007). *Formative and summative assessments in the classroom.* Retrieved from http://www.amle.org/Publications/WebExclusive/Assessment/tabid/1120/Default.aspx

Gibbs, J.C. (1991). Towards an Integration of Kohlberg's and Hoffman's Moral Development Theories. *Human Development, 34,* 88-104

Gibbs, J.C. (1993). Moral-cognitive interventions. In A.P. Goldstein & C.R. Huff (eds), *The Gang Intervention Handbook* (pp. 159-85). Champagin, IL: Research Press.

Gibbs, J.C. (2003). *Moral development and reality: Beyond the theories of Kohlberg and Hoffman.* Thousand Oaks, CA: Sage Publications

Gibbs, J.C., Basinger, K.S. & Fuller, D. (1992). *Moral Maturity: measuring the development of sociomoral reflection.* New Jersey: Lawrence Erlbaum Associates.

Gibbs, J.C., Basinger, K.S., Grime, R.L. & Snarey, J.R. (2007). Moral judgment development across cultures: Revisiting Kohlberg's universality claims. *Developmental Review, 27,* 443-500

Gibbs, J.C., Potter, G.B. & Goldstein, A.P. (1995). *The EQUIP program: Teaching youth to think and act responsibly through a peer-helping approach.* Champaign, IL, US: Research Press

Gipps, C.V. (1994). *Beyond Testing: Towards a Theory of Educational Assessment.* London: Falmer Press

Hatakka, M., Keskinen, E., Gregersen, N.P., Glad, A. & Hernetkoski, K. (2002). From control of the vehicle to personal self-control; broadening the perspectives to driver education. *Transportation Research,* Part F, 201-215

Helman, S., Grayson, G.B. & Parkes, A.M. (2010). *How can we produce safer new drivers?* Berkshire: TRL

Hoffman, M.L. (2000). *Empathy and moral development: Implications for caring and justice.* New York: Cambridge University Press

Infonu (2014). *Ergernissen in het verkeer.* Retrieved February 6, 2014, from http://auto-en-vervoer.infonu.nl/verkeer/5356-verkeer-ergernissen-in-het-verkeer-top-10.html

Jonah, B.A. (1997). Sensation seeking and risky driving: a review and synthesis of the literature. *Accident Analysis & Prevention, 29* (5), 651-665

Kane, M.T. (1992). An argument-based approach to validity. *Psychological Bulletin, 112,* 527-535

Kane, M.T. (2006). Validation. In R.L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17-64). Westport, CT: American Council on Education/Praeger.

Ker, K., Roberts, I., Collier, T., Beyer, F., Bunn, F. & Frost, C. (2004). Post-licence driver education for the prevention of road traffic crashes: a systematic review of randomised controlled trials. *Accident Analysis & Prevention, 37* (2), 305-313

Kettwich, C., Stocky, S., & Lemmer, U. (2010). Visual Behaviour of Car Drivers in Road Traffic. In L. Dorn (Eds.), *Driver Behavior and Training*. Farnham: Ashgate Publishing Limited (pp. 307-317

Kingston, N. & Nash, B. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice, 30* (4), 28-37

Kohlberg, L. (1983). *The philosophy of moral development: Moral stages and the idea of justice.* New York: Harper and Row.

Krantz, J. H., & Dalal, R. (2000). Validity of Web-based psychological research. In M. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 35–60). San Diego, CA: Academic Press.

Kuhlemeier, H., Boxtel, H. van & Til, A. van (2011). *Balans van de sociale opbrengsten in het basisonderwijs. Eerste meting voorjaar 2011.* Arnhem: Cito (PPON-reeks 48)

Kuiper, H., Penning, H. L. H. de, & Roelofs, E. C. (2011). *Toetsing in Simulatoren Fase 2*. Den Haag: TNO.

Leenders, I. & Brugman, D. (2005). The moral/non-moral domain shift in relation to delinquent behaviour in adolescents. *British Journal of Developmental Psychology, 23,* 65-79

Linden, W. van der (1982). A latent trait method for determining intrajudge inconsistency in the Angoff and Nedelsky techniques of standard setting, *Journal of Educational Measurement, 19* (4), 295-308

Mislevy, R.J. & Haertel, G.D. (2006). Implications of Evidence-Centered Design for Educational Testing. *Educational Measurement: Issues and Practice, 25* (4), 6-20

Mislevy, R.J., Steinberg, L.S., Almond, R.G. & Lukas, J.F. (2006). *Concepts, Terminology, and Basic Models of Evidence-Centered Design.* New Jersey: Lawrence Erlbaum Associates, Inc., Publishers

Moore, M. & Dahlen, E.R. (2008). Forgiveness and consideration of future consequence in aggressive driving. *Accident Analysis and Prevention, 40,* 1661-1666

Morrison, G.R., Ross, S.M. & Kemp, J.E. (2007). *Designing effective instruction* (5th edition), NY: John Wiley & Sons

Näätänen, R. & Summala, H. (1975). A simple method for simulating danger-related aspects of behaviour in hazardous activities. *Accident Analysis and Prevention, 7,* 63-70.

Nas, C.N., Brugman, D. & Koops, W. (2005). Effects of the EQUIP programme on the moral judgment, cognitive distortions, and social skills of juvenile delinquents. *Psychology, Crime, and Law, 11,* 421-434

Nicol, D.J. & MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice, *Studies in Higher Education, 31* (2), 199-218

Rakotonirainy, A., & Maire, F.D. (2005). *Context-Aware Driving Behavioural Model*. Paper presented on the 19th International Technical Conference on the Enhanced Safety of Vehicles, Washington DC

Reynolds, S.J. & Ceranic, T.L. (2007). The effects of Moral Judgment and Moral Identity on Moral Behavior: An Empirical Examination of the Moral Individual, *Journal of Applied Psychology, 92* (6), 1610-1624. DOI: 10.1037/0021-9010.92.6.1610

Roelofs, E. (2013). *Naar een ontwikkelingsgerichte opleiding.* Arnhem: Cito

Roelofs, E.C. (in preparation). *Modelgestuurde ontwikkeling van toetsopgaven.* Arnhem: Cito

Roelofs, E. & Sanders, P. (2007). Towards a Framework for Assessing Teacher Competence. *European Journal for Vocational Training, 40* (1), 123-139

Roelofs, E., Vissers, J. & Onna, M. van (2013). *A coaching program for recently licensed young drivers in the Netherlands: which drivers are attracted?* Paper presented on the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Bolton Landing

Roelofs, E., & Vissers, J. (2010). *CitoDrive: Verkeerseducatie voor Rijbewijsbezitters met Behulp van Instrumenten voor Verkeersassessment*. Paper presented on Nationaal Verkeersveiligheidscongres, Rotterdam.

Ryan, G.A., Legge, M., Diana, R. (1999). Age related changes in drivers' crash risk and crash type. *Accident Analysis & Prevention, 30* (3), 379-387

Schwebel, D.C., Severson, J., Ball, K.K. & Rizzo, M. (2006). Individual difference factors in risky driving: the role of anger/hostility, conscientiousness, and sensation seeking. *Accident Analysis and Prevention, 38,* 801-810

Siegrist, M. (1999). A causal model explaining the perception and acceptance of gene technology. *Journal of Applied Social Psychology, 29,* 2093-2106

Stams, G.J.J.M., Brugman, D., Dekovic, M., Rosmalen, L., van, Laan, P.L.C., van der & Gibbs, J.C. (2006). The moral judgment of juvenile delinquents: A meta-analyses. *Journal of Abnormal Child Psychology, 34,* 697-713

Stegge, H. (2006). Emotie en gewetensvorming. In: L. le Sage, H. Stegge & J. Sleutel (Red.), *Jeugddelinquentie en gewetensontwikkeling. Conceptualisering, diagnostiek en behandeling* (pp. 33-51). Amsterdam: Uitgeverij SWP.

Svenson, O. (1981). Are we all less risky and more skilful than our fellow drivers? *Acta Psychologica, 47,* 143-148.

Taras, M. (2005). Assessment – Summative and Formative – Some theoretical reflections. *British Journal of Educational Studies, 53* (4), 466-478

Treviño, L. T., Weaver, G., & Reynolds, S. J. (2006). Behavioral ethics in organizations: A review. *Journal of Management, 32,* 951–990.

Vissers, J.A.M.M., Betuw, A.M. van, Nägele, R.C., Kooistra, A. & Harteveld, M. (2007). *Learning goals for lifelong road safety education.* Amersfoort: DHV.

Vitaro, F., Brendgen, M. & Tremblay, R.E. (2000). Influence of deviant friends on delinquency: Searching for moderator variables. *Journal of Abnormal Child Psychology, 28* (4), 313-325

Vlakveld, W. (2006). *What does educational psychology tells us and what are the implications for simulator training? Proceedings of the Conference on European guidelines for the application of new technologies for driver training and education*. Humanist NoE: Madrid

Yagil, D. (1999). Gender and age-related differences in attitudes toward traffic laws and traffic violations. *Transportation Research Part F: Traffic Psychology and Behaviour, 1* (2), 123-135

Appendices

Appendix 1 – Assessment 1: Traffic Situations

Test 1 - Verkeerssituaties: jouw ervaringen en meningen (auto)

Deel 2

Hieronder volgen 43 stellingen.
Geef bij elke stelling aan hoe jij daarover denkt.

Je kan elke keer kiezen uit zes antwoordmogelijkheden, lopend van 'helemaal niet mee eens' tot 'helemaal mee eens'.

***1. Snelheidsregels zijn bedoeld voor anderen en niet voor mij**

Helemaal niet mee eens | Niet mee eens | Beetje oneens | Beetje mee eens | Mee eens | Helemaal mee eens

***2. Wat ik in het verkeer belangrijk vind, is dat ik kan rijden zoals ik wil**

Helemaal niet mee eens | Niet mee eens | Beetje oneens | Beetje mee eens | Mee eens | Helemaal mee eens

***3. Als ik vlak achter een ander rijd, komt dat doordat de andere bestuurder niet opschiet**

Helemaal niet mee eens | Niet mee eens | Beetje oneens | Beetje mee eens | Mee eens | Helemaal mee eens

***4. Sommige weggebruikers rijden zo slecht, dat ik wel moet claxonneren**

Helemaal niet mee eens | Niet mee eens | Beetje oneens | Beetje mee eens | Mee eens | Helemaal mee eens

***5. Als anderen zich niet aan de verkeersregels houden, dan hoef ik dat ook niet te doen**

Helemaal niet mee eens | Niet mee eens | Beetje oneens | Beetje mee eens | Mee eens | Helemaal mee eens

***6. Hoe ik ook rijd, er zijn altijd medeweggebruikers die er wat over te klagen hebben**

Helemaal niet mee eens | Niet mee eens | Beetje oneens | Beetje mee eens | Mee eens | Helemaal mee eens

***7. Ik kan veel hebben van weggebruikers die ik goed ken**

Helemaal niet mee eens | Niet mee eens | Beetje oneens | Beetje mee eens | Mee eens | Helemaal mee eens

***8. Als ik wil dat iemand opschiet, ga ik dicht achter diegene rijden**

Helemaal niet mee eens | Niet mee eens | Beetje oneens | Beetje mee eens | Mee eens | Helemaal mee eens

Print screen - page 1

Appendix 2 – Assessment 2: Self-evaluation of Driving Style

Test 2a - Zelfevaluatie rijgedrag (auto)

Situatie 1 van 18

***1. Je kijkt door de voorruit van je auto. Je nadert op ongeveer gelijke afstand als auto A de wegversmalling. Wat doe je?**

	00	10	20	30	40	50	60	70	80	90	100
Ik geef gas bij om er als eerste te zijn	○	○	○	○	○	○	○	○	○	○	○
Ik rem af om de ander voor te laten	○	○	○	○	○	○	○	○	○	○	○
Ik ga al eerder stil staan dan nu het geval is	○	○	○	○	○	○	○	○	○	○	○

Print screen – Item 1

Appendix 3 – Assessment 3: Motives for Driving Style

Test 3b - Motieven voor rijgedrag (auto)

Situatie 1 van 15

Dit was het voorbeeld.
Nu volgen er 15 situaties, met bijbehorende vragen.

Je nadert een gelijkwaardige kruising. Van rechts zie je fietsers aan komen rijden.

***1. Van de 10 keer dat dit voorkomt, hoe vaak rijd je door zonder voorrang te verlenen?**

0 1 2 3 4 5 6 7 8 9 10

Print screen – Item 1, part 1

Print screen – Item 1, part 2

Print screen – Item 1, part 3