How science informs engineering, education, and enforcement
A message for driving instructors

de Winter, J.C.F.; Kovacsova, Natalia

Publication date
2016

Document Version
Final published version

Published in
Handbook of Teen and Novice Drivers

Citation (APA)
de Winter, J. C. F., & Kovacsova, N. (2016). How science informs engineering, education, and enforcement: A message for driving instructors. In D. L. Fisher, J. K. Caird, W. J. Horrey, & L. M. Trick (Eds.), *Handbook of Teen and Novice Drivers: Research, Practice, Policy, and Directions* (pp. 31-45). Taylor & Francis.

4

How Science Informs Engineering, Education, and Enforcement: A Message for Driving Instructors

Joost C. F. de Winter

Natália Kovácsová

Abstract

The aim of this chapter is to illustrate to driving instructors how science contributes to cumulative knowledge on road safety. We do this by reviewing a scientific study for each of the three classical *E*s of road safety: (1) education, (2) enforcement, and (3) engineering.

Regarding education, we review the DeKalb experiment from the 1980s, which was a large-sample randomized controlled trial that studied the effect of driver education on postlicense crash rates. The DeKalb experiment showed that participants who were assigned to a state-of-the-art driver education program performed better on theory and road tests, and became licensed sooner than control participants who did not receive formal driving instruction. Although the state-of-the-art education improved these target outcomes, there is no consistent evidence that it reduced crash risk. The recent consensus is that theoretical knowledge and skillful maneuvering alone are

not sufficient for safe driving. Drivers should also have postlicense on-road experience and the lifestyle and attitudes that contribute to a safe driving style.

Regarding enforcement, we describe a UK study from the late 1990s on the statistical reliability of the formal road test. In this study, driving test candidates were asked to retake the test with a different examiner. The results showed surprisingly low consistency between the two tests, indicating that an assessment of a 30-minute drive might not be trustworthy. We provide several recommendations (such as increasing the test duration and implementing standardized routes and checklists) for improving the reliability of road testing. Furthermore, the value of computerized testing (e.g., hazard perception testing) and long-term data collection (e.g., in-vehicle driver state monitoring) is addressed.

Regarding engineering, the growing prevalence of active safety systems in vehicles has raised the question of how to treat such technologies in driver education curricula. A study on electronic stability control (ESC) was reviewed to illustrate how advances in technology improve road safety and affect elements of on-road training. In the case of ESC, skid training has become less relevant, but it is unknown whether learner drivers should experience critical driving situations during which the ESC gets activated. This may foster their overconfidence.

4.1 Introduction

Worldwide, 1.3 million fatal road traffic crashes occur on a yearly basis, making road injuries the eighth leading cause of death (Lozano et al., 2013). Young drivers are overrepresented, with 20–30% of the traffic fatalities resulting from crashes involving a driver under the age of 25 (Organisation for Economic Co-operation and Development [OECD], 2006). Fortunately, the high-income countries are making great strides in improving road safety (for more information, see Chapter 23). The ongoing implementation of road safety measures allows the setting of strict safety targets, with the long-term goal of zero fatalities in traffic (Rosencrantz et al., 2007).

4.1.1 The Three *E*s: Education, Enforcement, and Engineering

Road safety measures are traditionally categorized into the three *E*s: education, enforcement, and engineering (Learoyd, 1950; McKenna, 2012; Rothengatter, 1982). We define education as those mechanisms that intend to improve the knowledge and behavior of road users. This includes on-road practice, classroom courses, and mass media road safety campaigns (Beanland et al., 20113; Wakefield et al., 2010). Emerging methods such as simulator-based training (e.g., De Winter et al., 2009; Park et al., 2015) and in-vehicle monitoring systems that allow for real-time or postdrive feedback (e.g., Musicant & Lampel, 2010) also belong to the category of education (for further information, turn to Chapters 18 and 20). In North America and Australia, the term *driver education* is often used in reference to formal in-class and in-vehicle training prior to licensed driving (e.g., Mayhew & Simpson, 2002). Thus, driver education encompasses, and has a broader meaning than, driver training (see also Beanland et al., 2013). However, McKenna (2010) argued that in practice, people do not recognize the difference between the words *training* and *education*. In the present chapter, we use the term *education* for both classroom teaching and on-road instruction.

Enforcement includes the development and application of laws and regulations that aim to eliminate undesirable behaviors. Enforcement concerns not only such salient measures as police patrolling and speed cameras, but also driver testing, restricted driving in graduated driver licensing, breath alcohol testing, traffic regulations, vehicle safety standards and regulations, and laws regarding road design (Groeger & Banks, 2007; Zaal, 1994).

Engineering refers to the invention, design, construction, and modification of physical systems. Examples are modifications in road design such as black-spot treatments and traffic calming measures

(Elvik et al., 2009); the introduction of passive safety systems such as airbags and crumple zones; and, more recently, the introduction of active safety systems such as driver assistance and automation technology (e.g., Lee, 2007).

4.1.2 The Three *E*s and Driving Instructors

Among the three *E*s, driving instructors are probably most familiar with the first *E*, education. It is important that instructors know the scientific consensus and apply evidence-based education, not unlike clinicians who practice evidence-based medicine. However, education cannot be understood in isolation from the other two *E*s. After all, drivers drive in *engineered* vehicles and have to pass a formal driving test before being allowed to drive independently. Another example of the interaction between the three *E*s concerns the safety effectiveness of seat belts. Research has shown that the mere legislation of this technology in the 1970s (mandating that seat belts are installed in new cars and that it is compulsory to wear them) had limited effectiveness. It required substantial further investments in publicity campaigns and enforcement to ensure that people actually started wearing seat belts (Jonah et al., 1982; Mäkinen & Hagenzieker, 1991; Williams & Wells, 2004). Thus, driving instructors need to be familiar not only with the science behind education but also with issues of enforcement and engineering.

4.1.3 Aim of This Chapter

The aim of this chapter is to illustrate to driving instructors and other practitioners how the scientific method contributes to the development of road safety knowledge. We do this by describing three example scientific studies, one each in the areas of education (Stock et al., 1983), enforcement (Baughan & Simpson, 1999), and engineering (Farmer, 2006). For each of the three studies, we show the main results and explain the relevance for driving instructors. Furthermore, we discuss the limitations of these studies in an attempt to shed light on the limits of the acquired knowledge.

4.2 Education: Why Driver Education Sometimes Fails to Reduce Crashes

4.2.1 Prelicense Driver Education

One of the measures aiming to reduce novice driver crashes is prelicense driver education. The assumption that driver education produces safe drivers led to the introduction of formal driver education as a part of the licensing process in the first half of the twentieth century. The popularity of driver education grew in the 1950s and 1960s, stimulated by evaluation studies reporting that driver education was effective in reducing novice drivers' crash risk (see Mayhew, 2007 for a review). However, most of the early studies suffered from serious methodological weaknesses (e.g., no randomized controlled designs, small sample sizes), which means that the validity of their results is questionable.

4.2.2 Evaluation of Driver Education Effectiveness: The DeKalb Study

As a response to the growing popularity of driver education but ongoing concerns about its effectiveness, the National Highway Traffic Safety Administration (NHTSA) designed a state-of-art education program and a corresponding experiment to determine the effect of this program on road safety (Stock et al., 1983). This study took place between December 1977 and June 1981 in DeKalb County, Georgia. Herein, we report the results of the NHTSA final report (Stock et al., 1983) and reanalyses conducted by the Insurance Institute for Highway Safety (Lund et al., 1986) and by R. C. Peck & Associates (Peck, 2011).

We selected the DeKalb study as an illustration of a well-designed experiment. It had a large sample size and used a stratified randomization procedure for assigning participants to groups. Random assignment is considered to be a gold standard for investigating cause–effect relationships by ensuring that each participant has an equal chance of being placed in any group. Thus, at the end of the study, differences between groups can confidently be attributed to the effects of the experimental treatment (i.e., the type of driver education) on the dependent variables (i.e., indices of the effect of training). Another strength of the DeKalb study was that it evaluated educational effectiveness on measures of actual safety (i.e., crash and violation records from the Georgia Department of Administrative Services).

Students who had reached the age of 15 years (i.e., the legal licensing age), who did not already have a driver's license, who were not already participating in driver education, and who were motivated to obtain their driver's license as soon as possible could apply to participate in the DeKalb study (Stock et al., 1983). Over 16,000 secondary school students were randomly assigned to either of two educational groups or one control group while they were matched for sex, socioeconomic status, and grade point average (GPA). Students assigned to the first educational group participated in an advanced driver education program called the Safe Performance Curriculum (SPC). The SPC "was developed in such a way that it represented the best that the driver education community and its supporting scientific and technical resources had to offer as an accident countermeasure" (Riley & McBride, 1974, p. 5). Specifically, the SPC group received about 70 hours of formal education, consisting of three modes of formal instruction: (1) classroom instruction, including film-based driving simulation instruction, (2) instruction on a driving range, focusing on the initial development of vehicle control skills, skills in interacting with various roadway configurations, and emergency skills, and (3) on-road training focusing on the enhancement of the skills required in actual traffic. These types of formal instruction were complemented by practice-with-parents sessions and by guided learning designed to respond to individual needs (Riley & McBride, 1974; Weaver, 1978). In guided learning, the students could interact with an instructor during waiting intervals (e.g., when another group of students received the film-based driving simulation instruction). The duration of the in-vehicle instruction provided by the DeKalb study (range instruction and on-road training) was approximately one-third of the total time of formal instruction, whereas the remaining two-thirds was devoted to in-class education. The second group received a 20-hour education, which was called the predriver licensing curriculum (PDL) (Riley & McBride, 1974; Stock et al., 1983). The PDL aimed to develop only those skills and knowledge necessary for passing the driving test and covered less safety content. For example, the modules on hazard perception, alcohol and drugs, and skid control were not treated in the PDL. The control group did not receive any education provided by the DeKalb study (Stock et al., 1983; Weaver, 1978). It was expected that students assigned to this group were taught to drive by their parents or friends, or in commercial driving schools.

The results reported by Stock et al. (1983) were as follows:

- *Crashes and violations per assigned student*. About one year after the completion of the project, there were no statistically significant differences in the number of violations and crashes per student between the educational groups and the control group (Table 4.1 and Figure 4.1).
- *Crashes and violations per student who completed the course and obtained a driver's license*. During the first 6 months of licensed driving, there were slightly fewer crashes (average of 0.1021 [n = 3545], 0.1010 [n = 3375], and 0.1221 [n = 4135] for the SPC, PDL, and control groups, respectively) and violations (average of 0.1391, 0.1425, and 0.1753, respectively) for students in the SPC and PDL groups than for students in the control group, when analyzing only those students who had completed the SPC/PDL course and subsequently became licensed. These results are in line with the work of Peck (2011), who similarly concluded that the DeKalb study showed evidence of a small short-term crash and violation reduction *per licensed driver*. However, one limitation of these statistics is that not all assigned students actually completed the SPC/PDL course. The possibility that the more motivated/competent students completed the course, and hence skewed the results, cannot be ruled out. Stock et al. (1983) explained that "the percent of high GPA students among

TABLE 4.1 Crashes and Violations of All Assigned Students

		Crashes		Violations	
	Number of Assigned Students	% of Students with at Least One Crash	Mean Crashes per Student	% of Students with at Least One Violation	Mean Violations per Student
SPC	5464	28.61	.3776	45.59	.9771
PDL	5430	26.46	.3611	44.51	.9565
Control	5444	26.75	.3643	43.37	.9772

Source: Data from Tables II-7, II-8 and II-12 in Stock, J. R. et al., *Evaluation of Safe Performance Secondary School Driver Education Curriculum Demonstration Project* (Final Report DOT-HS-6-01462), National Highway Traffic Safety Administration, Washington, DC, 1983.

Note: PDL, predriver licensing; SPC, Safe Performance Curriculum. The crash and violation data were current as of December 1981 and December 1982, respectively.

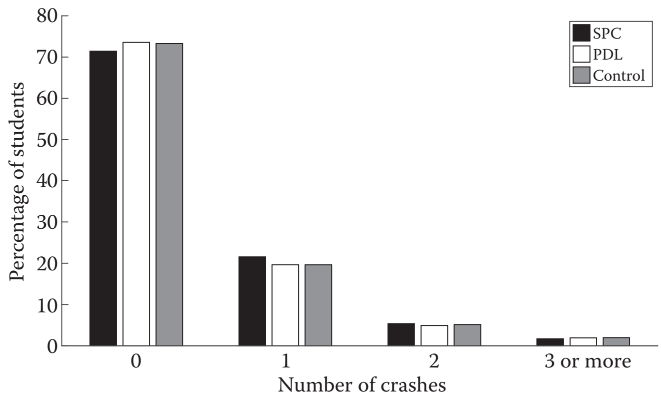

FIGURE 4.1 Distribution of the number of crashes for all assigned students (n = 5464 for Safe Performance Curriculum [SPC], n = 5430 for predriver licensing [PDL], n = 5444 for the control group). The crash and violation data were current as of December 1981 and December 1982, respectively. (Data from Table C-1 in Stock, J. R. et al., *Evaluation of Safe Performance Secondary School Driver Education Curriculum Demonstration Project* [Final Report DOT-HS-6-01462], National Highway Traffic Safety Administration, Washington, DC, 1983.)

the SPC group, 65.3 percent high GPA, and the PDL group, 65.8 percent high GPA, is somewhat higher than among the Control group, 59.6 percent high GPA. This difference probably reflects a self-selection factor in completing the SPC and PDL programs" (p. II-19).

- *Licensing rates*. Students assigned to SPC and PDL groups became licensed at greater rates compared to students assigned to the control group. Specifically, 70.6%, 66.7%, and 58.8% of students assigned to the SPC, PDL, and control groups, respectively, were licensed within 6 months of course completion or their 16th birthday, whichever was later.
- *Driving tests*. A subset of students completed additional tests of driving knowledge and skills. The SPC students scored higher than PDL students on a 56-item driving knowledge test administered on the last day of the quarter in which the student took driver education (the mean scores were 48.18 [n = 955] and 44.43 [n = 994], respectively). Furthermore, SPC students scored higher than the PDL and control groups on a standardized 30-minute on-road performance test, which was administered after the students were already licensed (mean percentages of correct behaviors were 68.75% [n = 100], 64.82% [n = 117], and 62.10% [n = 242], respectively).
- *Mileage*. By means of telephone surveys, it was determined that students in the control group had a higher driving exposure per licensed driver (the mean miles driven the day before the survey were 21.05 [n = 500] for SPC, 22.82 [n = 517] for PDL, and 24.93 [n = 498] for the control group, excluding 73, 73, and 80 students who reported they did not drive the previous day, respectively).

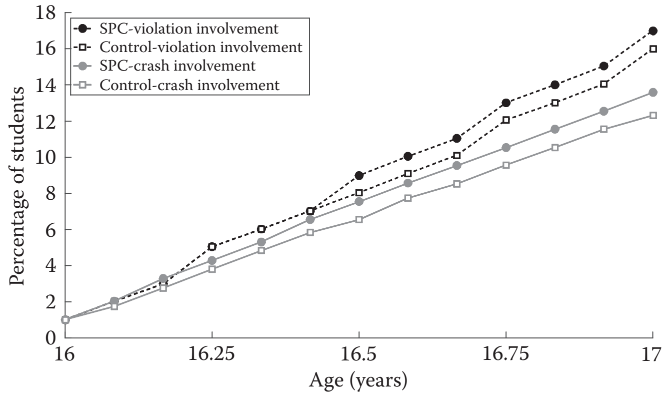

FIGURE 4.2 Estimated percentage of students having received a traffic violation and estimated percentage of students involved in a crash, per age group. The PDL data were omitted for clarity. (Data extracted from Figures 2 and 3 in Lund, A. K. et al., *Accid. Anal. Prev.*, 18, 349–357, 1986.)

Lund et al. (1986) reanalyzed the DeKalb data and applied a statistical model that controlled for students' GPA, parental education, parental occupation, sex, and the period during which they received the education. According to the statistical model by Lund et al. (1986), students assigned to the SPC group were 16% more likely to be licensed than students assigned to the control group. Furthermore, SPC students were 11% more likely to have crashed and 8% more likely to have received a traffic violation than the control group (see Figure 4.2).

4.2.3 Implications for Driving Instructors

The results of the DeKalb experiment yielded no consistent evidence that SPC and PDL programs reduced crash risk. A small crash reduction was observed per licensed driver, an effect that was detectable up to 18 months after licensure (Peck, 2011). However, if one wishes to express the effectiveness of the DeKalb study from a public-health per-capita point of view, the inescapable conclusion is that the "state-of-the-art" SPC program increased the likelihood of crashing compared to the control group. Whether one should adopt the per-licensed-driver (only students who completed the SPC course) or the per-capita (all students assigned to the course whether they completed it or not) perspective remains debatable (e.g., Peck, 2011).

One potential cause behind the limited safety effectiveness of the DeKalb program is that the SPC group in particular focused extensively on maneuvering at the driving range and on classroom instruction. Students in the SPC group indeed performed significantly better than the other two groups in a road test and a theory test. However, basic driving skills and knowledge about traffic rules are not sufficient for safe driving. The recent consensus is that drivers should acquire at least several months of independent postlicense driving experience in order to be safe drivers (Foss, 2011; Maycock & Lockwood, 1993). Appropriate lifestyle, attitudes, and skills for self-control are important prerequisites for safe driving as well (Hatakka et al., 2002; Jessor, 1987). In other words, although drivers clearly become more skillful and safe simply through learning by doing, risky driving attitudes are resistant to change. There is evidence that deliberate traffic violations, such as drunk driving, even *increase* with licensure (De Winter et al., 2015; Foss, 2011).

The DeKalb study demonstrated that students assigned to a driver education program became licensed sooner than students assigned to the control group. Although this is a positive outcome, it also means that the educational programs stimulated getting young people onto the roads who otherwise would not be driving, hence increasing the overall risk exposure. An additional issue is that young persons, males in particular, have riskier driving styles than older persons, due to their neurobiological immaturity (Dahl, 2008; Evans, 2006; Steinberg, 2008). To address these risks, current licensing systems aim to increase the licensing age and the amount of on-road driving experience prior to solo driving (OECD, 2006). For example, graduate licensing systems and multiphase driver educational programs worldwide aim to decrease fatalities by increasing the time period for achieving a full license and by letting novice drivers practice in protective conditions (OECD, 2006; Waller, 2003; Williams et al., 2012).

In the past decades, the effectiveness of driver education has been investigated in a number of studies (see Beanland et al., 2013 and Kardamanidis et al., 2010 for reviews on car driver education and motorcycle riding education, respectively). Unfortunately, many of these studies suffered from methodological weaknesses, such as attrition bias and a lack of randomized assignment (Beanland et al., 2013; Kardamanidis et al., 2010). Nonetheless, the available high-quality research indicates that driver education is useful for becoming skillful at the tasks that are the actual focus of the education. Examples of such target skills are to score highly on a road test, to perform well on a computerized test of a safety-relevant driving skill such as hazard anticipation, and to improve habits of wearing seat belts or helmets (e.g., Boele-Vos & De Craen, 2015; Horswill et al., 2015; Pradhan et al., 2009; Underwood et al., 2011). For example, in one randomized controlled trial in Thailand, it was found that driver education was successful in raising the proportion of motorcyclists who always wore helmets from 20.5% in the control group to 46.5% in the intervention group (Swaddiwudhipong et al., 1998). Furthermore, there is increasing evidence that safe driving skills can be acquired before licensure in simulator-based and PC-based training programs that target complex driving skills such as hazard anticipation, hazard mitigation, and attention, along with driver attitudes and motivation (Chapter 18; see also Chapters 21 and 28). In short, it would appear that drivers before licensure can develop important target skills that transfer to the open road, behaviors that are related to crash risk. However, the effect of education on actual crashes remains uncertain.

Despite the absence of consistent evidence that formal driver education reduces road traffic crashes, driver education continues to remain popular among the instructors who deliver it as well as among those who receive it (McKenna, 2012). In the last decades, driving instructors have rightly started to recognize that safe driving involves more than just theoretical knowledge of safe driving practices and skillful maneuvering at the driving range (Hatakka et al., 2002). It should be emphasized here that it is the research community that bears full responsibility for not having identified training programs that have been proven effective on actual measures of crash involvement. Driving instructors cannot be expected to develop and evaluate different training programs on their own. The driver education community is doing the very best it can with what researchers have given them as tools.

4.3 Enforcement: On the Statistical Reliability of On-Road Driver Testing

4.3.1 Driver Testing

In most countries, learner drivers have to pass a driving test in order to obtain their driver's license (Twisk & Stacey, 2007). Not only novice drivers but also professional and older persons with medical conditions have to participate in road tests (Siren & Haustein, 2015). Despite substantial advances in computerized visual and psychometric testing, the road test is still regarded as the gold standard of driver fitness (e.g., Dickerson et al., 2014; Rizzo et al., 2002). However, a study conducted in 1998 cast some doubt on the presumption that the outcome of the road test is particularly informative about the competence of a driver. Although road tests are closely tied to education, we treat driver testing as being in the category of enforcement because the driver's license indicates whether one is legally allowed to drive.

4.3.2 The Reliability of the Road Test

In November and December 1998, a study was undertaken at 20 test centers in the United Kingdom (Baughan & Simpson, 1999). Test candidates were asked whether they would like to take a second driving

TABLE 4.2 Number of Candidates Who Passed and Failed the Driving Tests

	Second Test	
First Test	Pass	Fail
Pass	80	57
Fail	75	154

Source: Baughan, C., & Simpson, B., Consistency of Driving Performance at the Time of the L-Test, and Implications for Driver Testing, in G. B. Grayson (ed.), *Behavioural Research in Road Safety IX*, Crowthorne: Transport Research Laboratory, pp. 206–214, 1999.

Note: The pass rate in the first test was 37.4% ([80 + 57]/366). The pass rate in the second test was 42.3% ([80 + 75]/366). This slight improvement in pass rates could indicate a learning effect.

test a few days later free of charge. The candidates were given a pass certificate if they passed the first test, the second, or both. Neither the candidate nor the examiner of the second driving test were provided with feedback about how the candidate had done in the first driving test until after the candidate had completed the second. A total of 366 candidates took part in the study.

The results revealed low consistency between the two tests (Table 4.2). Only in 64% of the driving tests were the results of the first and second tests the same. When expressed as a correlation coefficient, the test–retest reliability was $r = 0.25$. This is a weak association, especially when considering that the two driving tests were conducted at the same test center, thereby not incorporating regional differences in test difficulty.

4.3.3 Implications for Driving Instructors

In order to understand the implications of low test–retest reliability for driving instructors, it is useful to analyze where disagreement between the two driving tests could have arisen. Four sources of unreliability can be identified. First, there is the issue of interrater reliability. That is, even if two examiners independently assess the same driving test, they do not necessarily assign the same rating to this test, because humans differ regarding their perceptions and valuations (e.g., Boele-Vos & De Craen, 2015). Second, the capacities of the examiners as well as the candidates vary across time, because of momentary distractions as well as fluctuations in alertness, fatigue, and emotion. Third, the traffic conditions vary from one driving test to the other. That is, whether a candidate makes a mistake during a driving test depends on the behavior of other vehicles, weather conditions, and the route driven. Fourth, as explained by Baughan and Simpson (1999), it is likely that learner drivers apply for the driving test only when they are just sufficiently competent to pass the test (see also Baughan et al., 2005). A very poor driver will probably not apply for the road test but will continue practicing to increase the likelihood of passing. Therefore, driving test candidates are probably a homogeneous group, and no strong reliabilities are to be expected. Among statisticians, this phenomenon is known as *restriction of range*, whereby the association between two traits cannot be strong if all people are very much alike (see Kirkegaard, 2015 for an intuitive online demonstration).

Several recommendations can be put forward to improve the reliability of driving tests. First, it is possible to make the driving test longer. In the Baughan and Simpson (1999) study, the test lasted 35 minutes. Making the driving test longer will increase the amount of data (e.g., assessments, faults) that are collected and hence will increase test–retest reliability (Baughan & Simpson, 1999). Reliability can also be improved by using highly standardized routes and checklists, and by retraining the examiners such that they apply more homogeneous norming.

Another solution is to use computerized testing, such as video-based hazard perception tests and simulator-based testing (e.g., Horswill et al., 2015; Vlakveld, 2014; Chapter 28). The major advantage of computerized testing is that objective scoring is possible and that exactly the same traffic situations can be offered to all test candidates, guaranteeing a higher reliability than road testing. The disadvantage of computerized testing is the issue of validity. For example, it is known that people underestimate distance in driving simulators (e.g., Saffarian et al., 2015) and drive faster than they normally do in a car (Boer et al., 2000; De Groot et al., 2011). In addition, simulators are known to induce simulator sickness in a portion of the population, which means that they probably cannot be used for testing sensitive groups such as older drivers (e.g., Carsten & Jamson, 2011; also see Chapter 25).

A final lesson learned from the study by Baughan and Simpson (1999) is the fact that the situations we encounter, and our judgments thereof, are poorly replicable. Schmidt and Hunter (1999) explained that "the human central nervous system contains considerable noise at any given moment. This 'neural noise' can, for example, cause a person to answer two semantically identical questions differently, because of misreading a single word, because of a stray worry that popped up, etc." (p. 193). In order to obtain a statistically reliable assessment, driver behavior has to be recorded across long periods, and the collected data have to be aggregated across multiple measurement instances. In the near future, lifelong assessment and learning may indeed become a possibility. For example, driver state monitoring devices could be used for providing real-time alerts on risky driving behaviors and to keep track of one's driving style in the long term (Lee et al., 2015; Musicant & Lampel, 2010; also see Chapters 18 and 20). Furthermore, with such technology, parents can monitor their children's driving behavior via the Internet (Farmer et al., 2010).

4.4 Engineering: Electronic Stability Control Reduces Single-Vehicle Crashes by 40%

4.4.1 What Is Electronic Stability Control (ESC)?

Electronic stability control (ESC) is an active safety technology that aims to prevent skidding. The ESC system continuously compares the desired state of the vehicle (determined from the steering wheel angle and wheel speeds) with its current state (determined from the yaw rate and lateral acceleration). When the ESC detects that the vehicle is not traveling in the direction that it should be, it automatically applies the brakes of the individual wheels. For example, if the ESC detects that the yaw rate is smaller than the target yaw rate (*understeer*), it can brake the inner rear wheel in order to generate a corrective yaw moment. The ESC typically operates in conjunction with the engine and drivetrain systems, and can have additional functionalities such as rollover mitigation (Liebemann et al., 2004). In normal driving conditions, the driver cannot notice the presence of the ESC, because it is continuously analyzing sensor data but not implementing any corrective action. Only when the tires approach the maximum forces they can generate, the ESC applies a corrective braking action, in which case the driver may notice that an intervention has taken place.

4.4.2 Evaluation of the Safety Effectiveness of ESC

ESC was first introduced in 1995 and is now required for all passenger cars manufactured after September 2011 for sale in the United States (NHTSA, 2007). In the European Union, ESC is required in all new car models manufactured after November 2011 and in all newly registered cars from November 2014 onward (European Parliament and the Council for the European Union, 2009). The adoption of ESC and the subsequent requirement by various federal regulatory agencies that it be included in all manufactured vehicles is a consequence of accumulated scientific evidence supporting its safety effectiveness. With extensive test-track (e.g., Breuer, 1998) and driving simulator (e.g., Papelis et al., 2010) experiments, it has been shown that ESC has the potential to reduce crashes, in particular, loss-of-control and rollover crashes. However, the decisive scientific evidence came from actual on-road crash statistics.

There have been at least a dozen scientific publications on the on-road safety effectiveness of ESC (see Høye, 2011 for a review). We selected the work by Farmer (2006) as an exemplar because this is a representative study that features a large sample size and a straightforward method. Specifically, Farmer (2006) collected information on all police-reported crashes from 10 states for the years 2001–2003. He then extracted the number of crashes across 41 vehicle models having ESC as standard equipment and compared it to the same 41 vehicle models without ESC (or with ESC as option).

A total of 867 single-vehicle crashes were observed among the 41 ESC-equipped vehicles, while 1477 single-vehicle crashes were expected assuming that ESC-equipped vehicles had the same crash risk per registered vehicle as vehicles without ESC. Thus, because of ESC, single-vehicle crashes were reduced by 41% (i.e., 100% * [1477 − 867]/1477). The calculation of the expected crash risk included a correction factor (between 2% and 8%) to account for vehicle age. This correction factor was applied because it is known that older vehicles are more likely to be involved in car crashes, for example, because the quality of the vehicle has deteriorated or because older vehicles are driven by people who adopt riskier driving styles (e.g., teen drivers driving second-hand cars).

Additionally, Farmer (2006) found that ESC reduced injury crashes by 45% (337 observed versus 617 expected crashes) and fatal crashes by 56% (89 observed versus 204 expected crashes). The safety gains of ESC were even greater for rollover crashes, where 39 crashes were observed and 163 expected, an impressive reduction of 76%. Based on these numbers, it is clear why ESC has been called "the greatest safety innovation since the safety belt" (Nason, 2006). The safety effectiveness of ESC is especially good news for male novice drivers, who are known to be overinvolved in single-vehicle crashes (Laapotti & Keskinen, 1998).

These promising results must be somewhat tempered because single-vehicle crashes accounted for only 12% of all police-reported crashes (Farmer, 2006). Because ESC is designed to prevent loss-of-control crashes, it is perhaps not surprising that ESC had no statistically significant effect on multiple-vehicle crashes (Farmer, 2006). Several other studies have found that ESC even slightly *increases* certain types of multiple-vehicle crashes, such as rear-end collisions (Høye, 2011). A possible explanation is a phenomenon called *behavioral adaptation*. When drivers know that ESC is present in their cars, they may feel more confident and adopt riskier driving styles (Kulmala & Rämä, 2013). On the other hand, self-selection and police-reporting bias cannot be ruled out. For example, ESC-equipped car crashes may be more likely to be entered into the police records for the simple reason that equipped cars are more expensive or used by different types of drivers than nonequipped cars (Scully & Newstead, 2008). This could mean that the crash reduction potential of ESC is actually underestimated.

4.4.3 Implications for Driving Instructors

The growing prevalence of ESC has clear implications for driver education. One evident example is skid training, which becomes less important as ESC becomes more prevalent (Barker & Woodcock, 2011). An important question is whether learner drivers should experience the functionality of ESC, for example, by means of a skid pad or high-speed cornering exercise. Although learning by experiencing seems a sensible thing to do, there are potential downsides. Letting learner drivers experience the limits of the vehicle may indeed improve their handling skills but could also lead to behavioral adaptation and overconfidence (e.g., Beanland et al., 2013; Katila et al., 1996; McKenna, 2012).

ESC as well as other types of technologies, such as route navigation devices, blind-spot monitors, and advanced emergency braking systems (AEBS), are gradually finding their way into consumer vehicles. Ongoing research is trying to determine how to treat such technologies in driver education curricula (Hedlund, 2007; Panou et al., 2010). It is currently possible for a student to be trained in a car with automatic transmission and to take the driving test in such a car (in which case, in some jurisdictions, the driver's license does not permit driving a vehicle with manual transmission). In the future, driver education and licensing procedures will have to be adjusted to include highly automated driving and the use of in-vehicle interfaces (see Hancock & Parasuraman, 1992 for an early discussion on this topic).

Of course, not all technology is beneficial for road safety. Cell phones and infotainment devices can seriously undermine safety, especially in teen and novice drivers who like to stay in contact with peers and have little spare mental capacity for performing secondary tasks (Lee, 2007; Young & Stanton, 2007; also see Chapter 12). It has been recommended that driver education should improve learner drivers' awareness of their risky habits (Hatakka et al., 2002).

4.5 Discussion and Conclusion

The aim of the present chapter is to illustrate to driving instructors and other stakeholders how science contributes to the expansion of knowledge on road safety. We provided three examples, one for each *E*: education (the DeKalb driver education study by Stock et al., 1983), enforcement (the study on the reliability of the road test by Baughan & Simpson, 1999), and engineering (the study on the effectiveness of ESC by Farmer, 2006). These examples provide an illustration of how research has contributed to cumulative knowledge.

The three selected papers rely on a number of scientific methods, such as a randomized controlled trial, where it is only the effect of the treatment, not some other factor, that can explain why the treatment produces whatever results are observed (Stock et al., 1983); the blinding of experimental conditions to the individuals involved in the evaluation so that bias the candidate or examiner might have is removed from the assessment (Baughan & Simpson, 1999); and systematic archiving and analysis of crash data (Farmer, 2006). In essence, these methods are intended to protect scientists from self-deception. This is important because humans all have certain ideas and conceptions of how the world works, and this may bias their observations. As explained by Wolpert (1994) in his book *The Unnatural Nature of Science*, "ordinary, day-to-day common sense—will never give an understanding about the nature of science" (p. xi).

Although the authors of the present handbook write about novice and teen drivers, they do not necessarily have firsthand experience in automotive engineering, police enforcement, or driver education. In fact, an author of a chapter in this handbook and a leading authority on the value of hazard perception testing in the licensing process openly admits he does not have a driver's license, and he had the following proposition in his PhD thesis: "It is an advantage to study driver behaviour without having a driving licence" (Vlakveld, 2011). Vlakveld's position is not strange or absurd. Considering the wide array of biases and predispositions toward driving (Vanderbilt, 2008), it seems reasonable that scientists—in their quest for objectivity—dissociate themselves from the activity of driving and devote their attention to science.

In this chapter, we showed several things: (1) driver education is known to improve target skills (e.g., obtaining a driver's license), but whether it actually reduces crashes compared to informal education remains unproven; (2) a subjective assessment of a 30-minute drive is statistically unreliable; and (3) ongoing technological innovations, including ESC, have a major positive impact on road safety. We argue that future driving will look different from today. Most likely, there will be more in-vehicle technologies, more automated driving systems, more data on driver and vehicle state, and more vehicle-to-vehicle and vehicle-to-infrastructure communication than exist today. These developments will allow us to predict, prevent, and mitigate crashes with ever-greater effectiveness. The need for driver education is not likely to disappear. It is true that automatically driving cars may one day be the norm. However, just as pilots need to interpret a large number of displays in the cockpit and to take over control when automation fails, so too will drivers need to know how to take over control when the automated driving suite fails or reaches its functional limitations. Thus, driver education may become even more critical with the emergence of technology.

Acknowledgments

This work has been supported in part by the Marie Curie Actions of the European Union's Seventh Framework Programme FP7/2007–2013 under REA grant agreement number 608092.

References

Barker, P., & Woodcock, A. (2011). Driving skills, education and in-vehicle technology. *International Journal of Vehicle Design, 55*, 189–207.

Baughan, C., & Simpson, B. (1999). Consistency of driving performance at the time of the L-test, and implications for driver testing. In G. B. Grayson (Ed.), *Behavioural Research in Road Safety IX* (pp. 206–214). Crowthorne: Transport Research Laboratory.

Baughan, C. J., Gregersen, N. P., Hendrix, M., & Keskinen, E. (2005). *Towards European Standards for Testing* (Final Report). Brussels: Commission Internationale des Examens de Conduite Automobile (CIECA). Retrieved from http://www.cieca.eu/sites/default/files/documents/projects_and_studies/EU_TEST_Project_Final_Report.pdf.

Beanland, V., Goode, N., Salmon, P. M., & Lenné, M. G. (2013). Is there a case for driver training? A review of the efficacy of pre- and post-licence driver training. *Safety Science, 51*, 127–137.

Boele-Vos, M. J., & De Craen, S. (2015). A randomized controlled evaluation study of the effects of a one-day advanced rider training course. *Accident Analysis & Prevention, 79*, 152–159.

Boer, E. R., Yamamura, T., Kuge, N., & Girshick, A. (2000). Experiencing the same road twice: A driver centered comparison between simulation and reality. *Proceedings of Driving Simulation Conference*, Paris, France, 33–55.

Breuer, J. J. (1998). Analysis of driver–vehicle interactions in an evasive manoeuvre—Results of "moose test" studies. *16th International Technical Conference on the Enhanced Safety of Vehicles*, Windsor, Ontario, 620–627. Retrieved from http://www-nrd.nhtsa.dot.gov/pdf/nrd-01/Esv/esv16/98S2W35.PDF.

Carsten, O., & Jamson, A. H. (2011). Driving simulators as research tools in traffic psychology. In B. E. Porter (Ed.), *Handbook of Traffic Psychology* (pp. 87–96). London: Academic Press.

Dahl, R. E. (2008). Biological, developmental, and neurobehavioral factors relevant to adolescent driving risks. *American Journal of Preventive Medicine, 35*, S278–S284.

De Groot, S., De Winter, J. C. F., Mulder, M., & Wieringa, P. A. (2011). Nonvestibular motion cueing in a fixed-base driving simulator: Effects on driver braking and cornering performance. *Presence: Teleoperators and Virtual Environments, 20*, 117–142.

De Winter, J. C. F., De Groot, S., Mulder, M., Wieringa, P. A., Dankelman, J., & Mulder, J. A. (2009). Relationships between driving simulator performance and driving test results. *Ergonomics, 52*, 137–153.

De Winter, J. C. F., Dodou, D., & Stanton, N. A. (2015). A quarter of a century of the DBQ: Some supplementary notes on its validity with regard to accidents. *Ergonomics*, *58*, 1745–1769.

Dickerson, A. E., Brown Meuel, D., Ridenour, C. D., & Cooper, K. (2014). Assessment tools predicting fitness to drive in older adults: A systematic review. *The American Journal of Occupational Therapy, 68*, 670–680.

Elvik, R., Høye, A., Vaa, T., & Sørensen, M. (2009). *The Handbook of Road Safety Measures*. Bingley: Emerald Group Publishing.

European Parliament and the Council for the European Union (2009). Regulation (EC) No 661/2009. Concerning type-approval requirements for the general safety of motor vehicles, their trailers and systems, components and separate technical units intended therefor. Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32009R0661.

Evans, L. (2006). Innate sex differences supported by untypical traffic fatalities. *Chance, 19*, 10–15.

Farmer, C. M. (2006). Effects of electronic stability control: An update. *Traffic Injury Prevention, 7*, 319–324.

Farmer, C. M., Kirley, B. B., & McCartt, A. T. (2010). Effects of in-vehicle monitoring on the driving behavior of teenagers. *Journal of Safety Research, 41*, 39–45.

Foss, R. D. (2011). *Measuring Changes in Teenage Driver Crash Characteristics during the Early Months of Driving*. Washington, DC: AAA Foundation for Traffic Safety.

Groeger, J. A., & Banks, A. P. (2007). Anticipating the content and circumstances of skill transfer: Unrealistic expectations of driver training and graduated licensing? *Ergonomics, 50*, 1250–1263.

Hancock, P. A., & Parasuraman, R. (1992). Human factors and safety in the design of intelligent vehicle–highway systems (IVHS). *Journal of Safety Research, 23*, 181–198.

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: Traffic Psychology and Behaviour, 5*, 201–215.

Hedlund, J. (2007). Novice teen driving: GDL and beyond. *Journal of Safety Research, 38*, 259–266.

Horswill, M. S., Hill, A., & Wetton, M. (2015). Can a video-based hazard perception test used for driver licensing predict crash involvement? *Accident Analysis & Prevention, 82*, 213–219.

Høye, A. (2011). The effects of Electronic Stability Control (ESC) on crashes—An update. *Accident Analysis & Prevention, 43*, 1148–1159.

Jessor, R. (1987). Risky driving and adolescent problem behavior: An extension of problem-behavior theory. *Alcohol, Drugs, & Driving, 3*, 1–11.

Jonah, B. A., Dawson, N. E., & Smith, G. A. (1982). Effects of a selective traffic enforcement program on seat belt usage. *Journal of Applied Psychology, 67*, 89–96.

Kardamanidis, K., Martiniuk, A., Ivers, R. Q., Stevenson, M. R., & Thistlethwaite, K. (2010). Motorcycle rider training for the prevention of road traffic crashes. *Cochrane Database of Systematic Reviews, 10*, CD005240.

Katila, A., Keskinen, E., & Hatakka, M. (1996). Conflicting goals of skid training. *Accident Analysis & Prevention, 28*, 785–789.

Kirkegaard, E. (2015). Understanding restriction of range. Retrieved from https://emilkirkegaard.shinyapps.io/Understanding_restriction_of_range.

Kulmala, R., & Rämä, P. (2013). Definition of behavioural adaptation. In C. M. Rudin-Brown, & S. L. Jamson (Eds.), *Behavioural Adaptation and Road Safety: Theory, Evidence and Action* (pp. 11–21). Boca Raton: CRC Press.

Laapotti, S., & Keskinen, E. (1998). Differences in fatal loss-of-control accidents between young male and female drivers. *Accident Analysis & Prevention, 30*, 435–442.

Learoyd, C. G. (1950). The carnage on the roads. *The Lancet, 255*, 367–369.

Lee, J. D. (2007). Technology and teen drivers. *Journal of Safety Research, 38*, 203–213.

Lee, Y.-C., Belwadi, A., Bonfiglio, D., Malm, L., & Tiedeken, M. (2015). Techniques for reducing speeding beyond licensure: Young drivers' preferences. *Proceedings of the Eighth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Salt Lake City, Utah, 168–174. Retrieved from http://drivingassessment.uiowa.edu/sites/default/files/DA2015/papers/026.pdf.

Liebemann, E. K., Meder, K., Schuh, J., & Nenninger, G. (2004). Safety and performance enhancement: The Bosch electronic stability control (ESP). *SAE Paper*, 2004-21-0060. Retrieved from http://www-nrd.nhtsa.dot.gov/pdf/nrd-01/esv/esv19/05-0471-O.pdf.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., & Cross, M. (2013). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet, 380*, 2095–2128.

Lund, A. K., Williams, A. F., & Zador, P. (1986). High school driver education: Further evaluation of the DeKalb County study. *Accident Analysis & Prevention, 18*, 349–357.

Mäkinen, T., & Hagenzieker, M. P. (1991). Strategies to increase the use of restraint systems—Background paper. *Proceedings of a workshop organized by SWOV and VTT at the VTI-TRB International Conference Traffic Safety on Two Continents*. Leidschendam: SWOV, 2–5.

Maycock, G., & Lockwood, C. R. (1993). The accident liability of British car drivers. *Transport Reviews, 13*, 231–245.

Mayhew, D. R. (2007). Driver education and graduated licensing in North America: Past, present, and future. *Journal of Safety Research, 38*, 229–235.

Mayhew, D. R., & Simpson, H. M. (2002). The safety value of driver education and training. *Injury Prevention, 8*, ii3–ii8.

McKenna, F. (2010). *Education in Road Safety: Are We Getting Right?* (Report No. 10/113). London: RAC Foundation.

McKenna, F. P. (2012). How should we think about the three E's: education, engineering and enforcement? *The 5th International Conference on Traffic and Transport Psychology*, Groningen, the Netherlands, August 29–31, 2012. Retrieved from http://www.icttp2012.com/images/stories/presentations/29th/Keynote/McKenna_ICTTP2012_Keynote_How_should_we_think.pdf.

Musicant, O., & Lampel, L. (2010). When technology tells novice drivers how to drive. *Transportation Research Record: The Journal of the Transportation Research Board, 2182*, 8–15.

Nason, N. R. (2006). Remarks of Nicole R. Nason, Administrator, National Highway Traffic Safety Administration U.S.A. Made at the *140th Session of the United Nations' World Forum for Harmonization of Vehicle Registration* (WP.29). Geneva, November.

NHTSA. (2007). 49 Code of Federal Regulations (CFR) Parts 571 and 585. US Government Printing Office, Washington, DC.

Organisation for Economic Co-operation and Development. (2006). *Young Drivers: The Road to Safety*. Paris: OECD.

Panou, M. C., Bekiaris, E. D., & Touliou, A. (2010). ADAS module in driving simulation for training young drivers. *13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 1582–1587.

Papelis, Y. E., Watson, G. S., & Brown, T. L. (2010). An empirical study of the effectiveness of electronic stability control system in reducing loss of vehicle control. *Accident Analysis & Prevention, 42*, 929–934.

Park, G. D., Allen, R. W., & Rosenthal, T. J. (2015). Novice driver simulation training potential for improving hazard perception and self-confidence while lowering speeding risk attitudes for young males. *Proceedings of the Eighth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 253–259. Retrieved from http://drivingassessment.uiowa.edu/sites/default/files/DA2015/papers/039.pdf.

Peck, R. C. (2011). Do driver training programs reduce crashes and traffic violations?—A critical examination of the literature. *IATSS Research, 34*, 63–71.

Pradhan, A. K., Pollatsek, A., Knodler, M., & Fisher, D. L. (2009). Can younger drivers be trained to scan for information that will reduce their risk in roadway traffic scenarios that are hard to identify as hazardous? *Ergonomics, 52*, 657–673.

Riley, M. C., & McBride, R. S. (1974). *Safe Performance Curriculum for Secondary School Driver Education: Program Development, Implementation, and Technical Findings* (Final Report. DOT-HS-003-2-427). Washington, DC: National Highway Traffic Safety Administration.

Rizzo, M., Jermeland, J., & Severson, J. (2002). Instrumented vehicles and driving simulators. *Gerontechnology, 1*, 291–296.

Rosencrantz, H., Edvardsson, K., & Hansson, S. O. (2007). Vision zero–Is it irrational? *Transportation Research Part A: Policy and Practice, 41*, 559–567.

Rothengatter, T. (1982). The effects of police surveillance and law enforcement on driver behaviour. *Current Psychological Reviews, 2*, 349–358.

Saffarian, M., De Winter, J. C. F., & Senders, J. W. (2015). Measuring drivers' visual information needs during braking: A simulator study using a screen-occlusion method. *Transportation Research Part F: Traffic Psychology and Behaviour, 33*, 48–65.

Schmidt, F. L., & Hunter, J. E. (1999). Theory testing and measurement error. *Intelligence, 27*, 183–198.

Scully, J., & Newstead, S. (2008). Evaluation of electronic stability control effectiveness in Australasia. *Accident Analysis & Prevention, 40*, 2050–2057.

Siren, A., & Haustein, S. (2015). Driving licences and medical screening in old age: Review of literature and European licensing policies. *Journal of Transport & Health, 2*, 68–78.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28*, 78–106.
Stock, J. R., Weaver, J. K., Ray, H. W., Brink, J. R., & Sadof, M. G. (1983). *Evaluation of Safe Performance Secondary School Driver Education Curriculum Demonstration Project* (Final Report DOT-HS-6-01462). Washington, DC: National Highway Traffic Safety Administration.
Swaddiwudhipong, W., Boonmak, C., Nguntra, P., & Mahasakpan, P. (1998). Effect of motorcycle rider education on changes in risk behaviours and motorcycle-related injuries in rural Thailand. *Tropical Medicine & International Health, 3*, 767–770.
Twisk, D. A. M., & Stacey, C. (2007). Trends in young driver risk and countermeasures in European countries. *Journal of Safety Research, 38*, 245–257.
Underwood, G., Crundall, D., & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*, 435–446.
Vanderbilt, T. (2008). *Traffic. Why We Drive the Way We Do (and What It Says about Us)*. London: Allen Lane.
Vlakveld, W. P. (2011). Hazard anticipation of young novice drivers: Assessing and enhancing the capabilities of young novice drivers to anticipate latent hazards in road and traffic situations. Doctoral Dissertation, University of Groningen, the Netherlands.
Vlakveld, W. P. (2014). A comparative study of two desktop hazard perception tasks suitable for mass testing in which scores are not based on response latencies. *Transportation Research Part F: Traffic Psychology and Behaviour, 22*, 218–231.
Wakefield, M. A., Loken, B., & Hornik, R. C. (2010). Use of mass media campaigns to change health behaviour. *The Lancet, 376*, 1261–1271.
Waller, P. F. (2003). The genesis of GDL. *Journal of Safety Research, 34*, 17–23.
Weaver, J. K. (1978). Does quality driver education produce safer drivers? *Traffic Safety, 78*, 18–20.
Williams, A. F., & Wells, J. K. (2004). The role of enforcement programs in increasing seat belt use. *Journal of Safety Research, 35*, 175–180.
Williams, A. F., Tefft, B. C., & Grabowski, J. G. (2012). Graduated driver licensing research, 2010–present. *Journal of Safety Research, 43*, 195–203.
Wolpert, L. (1994). *The Unnatural Nature of Science*. Cambridge: Harvard University Press.
Young, M. S., & Stanton, N. A. (2007). What's skill got to do with it? Vehicle automation and driver mental workload. *Ergonomics, 50*, 1324–1339.
Zaal, D. (1994). *Traffic Law Enforcement: A Review of the Literature* (Report No. 53). Leidschendam: SWOV.