TRANSPORTATION RESEARCH

CIRCULAR

Number E-C101 August 2006

Driver Education

The Path Ahead

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

TRANSPORTATION RESEARCH BOARD
2006 EXECUTIVE COMMITTEE OFFICERS

TRANSPORTATION RESEARCH BOARD
2006 TECHNICAL ACTIVITIES COUNCIL

TRANSPORTATION RESEARCH CIRCULAR E-C101

Driver Education

The Path Ahead

Transportation Research Board
Operator Education and Regulation Committee

August 2006

Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001
www.TRB.org

TRANSPORTATION RESEARCH CIRCULAR E-C101
ISSN 0097-8515

The **Transportation Research Board** is a division of the National Research Council, which serves as an independent adviser to the federal government on scientific and technical questions of national importance. The National Research Council, jointly administered by the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine, brings the resources of the entire scientific and technical communities to bear on national problems through its volunteer advisory committees.

The **Transportation Research Board** is distributing this Circular to make the information contained herein available for use by individual practitioners in state and local transportation agencies, researchers in academic institutions, and other members of the transportation research community. The information in this Circular was taken directly from the submission of the authors. This document is not a report of the National Research Council or of the National Academy of Sciences.

Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001
www.TRB.org

Jennifer Correro, Proofreader and Layout

Contents

Preface

This circular summarizes the proceedings of the midyear meeting and workshop conducted by the Transportation Research Board's (TRB's) Operator Education and Regulation Committee (ANB30), September 12–13, 2005, in Washington, D.C. The meeting was cochaired by Committee Chair Dan Mayhew and Dr. A. James McKnight. The committee thanks Dr. McKnight for planning, organizing, and conducting the workshop and for preparing this circular. The committee also thanks each of the authors who prepared and presented papers and then prepared the summaries appearing in this circular.

Introduction

A midyear meeting of TRB's Operator Education and Regulation Committee (ANB30) in 1998 presented a number of issues related to the topic of Driver Education at the Crossroads. In the years that have passed since the 1998 meeting, much has happened. A midyear meeting held in 2005, Driver Education: The Path Ahead, offered a series of papers that describe recent findings and developments.

In his paper, The Novice Driver Problem, David Preusser presents recent data showing the extremely high accident rate of novices, one that drops by two thirds in the first months of driving. This focuses the need for instruction on the period right after licensing, when beginner errors expose novices to risk. The challenge is to overcome the mistakes through instruction that is more informative and motivating than that which has prevailed in the past.

In Content of Driver Education, A. James McKnight explains how we can help reduce the high initial accident rate by focusing on the errors that are the biggest contributors. They include frequent lapses of attention, not looking where hazards lurk or recognizing them when they appear, and not adjusting speed to road and traffic conditions. Clearly, reckless, high-speed driving is negligible among novices.

In Instructional Methods for Young Drivers, John F. Brock explains that methods available today are far more interactive and individualized than traditional classroom instruction. Computer-based training is highly learner centered and, with the motion visual capabilities (e.g., DVD), can put students behind the wheel where they can be exposed to and learn to handle the many threats to safety before they are encountered on the road.

Student Competency Measures, according to author Larry Lonero, still depends primarily on written and road tests, but advances in simulation allow measures to include situations that occur too rarely or are too dangerous on the road to make a part of student assessment. There remains a need to set competency standards that will satisfy both licensing and safety needs.

In Novice Driver Training Effectiveness Evaluation, Raymond C. Peck observes that suitable alternatives must be found to accidents, which, through the ultimate criterion of safety, happen too infrequently to allow valid assessment without prohibitively large samples. The most valid alternative would be measures of the way novices drive, not how they can drive but the way they actually do drive when on their own. Effectiveness can only be measured through random experiments in which control groups receive sufficient instruction to become licensed and face the same challenges as those taking the driver education program being evaluated.

In his paper, The Future of Driver Education, Daniel R. Mayhew charts a course for driver education taking into account the findings presented in the above papers. The failure of present-day instruction to provide a convincing demonstration of its ability to reduce accidents establishes the need for change. Recent research revealing the magnitude and nature of novice driver shortcomings may lead to adjustment of objectives, content, and methods of instruction. Advances in technology that can bring the actual demands of driving into the educational process offer a means of helping novices to better recognize and cope with the demands of safety on the highway.

The Novice Driver Problem

DAVID F. PREUSSER
Preusser Research Group, Inc.

The rate of serious crash involvement for 16-year-olds is estimated at 34.5 per million miles driven; 20.2 for 17-year-olds; and 13.8 for 18-year-olds. This compares with 7.8 for drivers in their 20s and 3.9 for drivers ages 60 to 69. Clearly, crash rates for teens are much higher than comparable rates for older drivers. Major risk factors for teen drivers have been identified. These include lack of driving experience, licensing at young ages, passengers, night driving, and alcohol.

DRIVING EXPERIENCE

High teen driver crash rates have been related to lack of driving experience. Crash rates for teens and newly licensed drivers of other ages are extremely high during the first few hundred, or just a few thousand, miles or kilometers of driving. Crash rates drop quickly as the new driver gains experience. Crash rates eventually stabilize after tens of thousands of miles of driving.

A study of young Norwegian drivers indicated a sharp decline in crash rates per kilometer driven during the first few months of driving. Similar results have been reported in Sweden, Canada, and Michigan.

One recent study focused on the relationship between risk and experience during the first few hundred miles of driving for 16-year-olds following licensure. The results indicated an extremely high crash rate immediately following licensing and lasting for the first few months or 1,000 mi of driving. In fact, the crash rate for both males and females was estimated to be three times greater during the first 1,000 mi of driving as compared to the next 2,000 to 3,000 mi of driving.

YOUNG AGES

States that allow early licensing, at age 16 years and zero months or earlier, tend to have higher teen crash rates than states that delay licensing. Early licensing is promoted through allowing learning driving to begin at an early age, say 14 or 15. Time discounts for completing driver education (DE) can also contribute to earlier licensure.

- Night: Fatal crash rates for all age groups are higher at night than during the day. This difference is particularly great for 16- and 17-year-olds.
- Passengers: Teen drivers accompanied by passengers are twice as likely to be involved in a fatal crash as teen drivers traveling alone. Multiple teen passengers double crash risk again.
- Alcohol: The Age 21 drinking laws and Zero Tolerance have contributed to substantial reductions in the number of young persons who choose to drink and drive. Still, drinking and driving must be included in any list of teen driving crash risk factors.

All of the above factors should be addressed in a comprehensive approach to reducing teen crashes. DE will likely play a role in addressing the difficulties of dealing with night driving, the need to avoid distractions and remain focused on the driving task when with teen passengers, and the absolute necessity to avoid driving after taking any alcohol or drugs.

Future DE may also be able to reduce crashes by focusing on the initial “errors of inexperience” rather than trying to develop a “lifetime of responsible driving” within the context of a “thirty and six” program. Clearly, lack of driving experience is a major contributor to the high crash rate for young drivers and the extremely high crash rate during the first few months of driving. Research and development of future DE programs should consider the possibility of focusing on the problems faced by young drivers during this initial period.

Content of Driver Education

A. JAMES MCKNIGHT
Transportation Research Associates

The steep decline in accidents in the first months of driving as described in the previous presentation—two-thirds drop in the first 1,000 mi—evidences the effects of experience in leading to safer driving by novices. While maturity is a factor in reducing the accident rate of young drivers its impact is experienced over many years, not just the first few months of driving. The magnitude and timing of the decline in novice accidents suggests that the initially high rate is more likely attributable to the errors of inexperience than to the high speed and irresponsible behavior often associated with teen driving in the public's mind. This impression comes mainly from the fatalities that grab headlines but account for only 1% of all injury accidents involving teens. While the devastating effect of fatal accidents upon the families of teen drivers gives a high priority to means of combating them, DE is not a promising form of intervention. Films and other devices intended to discourage irresponsible behavior through scare tactics have only momentary effects and don't address the specific behavior leading to accidents. More productive routes to fatality reduction have been through enforcement, vehicle design, and occupant protection.

NOVICE DRIVING ERRORS

Focusing the content of DE upon those errors that account for the largest share of novice accidents offers the opportunity to make significant inroads upon the high initial accident rate. A study of accident reports shows the largest single category of error involves visual search, defined as not looking for the right things at the right time along the road ahead and to the side for cars and people who might enter the path, or behind when slowing, backing, or changing lanes. Next are attention errors, with the eyes pointed in the right direction but the mind somewhere else. Often it is the result of distraction or having to share attention among two or more situations. The third major category involves speed, primarily not adjusting adequately to traffic or to curves and slick surfaces. Very high speeds were involved in less than 1% of reported injury accidents. Less frequent yet still important errors involve maintaining space between vehicles when following or crossing the paths of others; handling emergencies by swerving to avoid collisions and controlling skids; basic control of the car; following rules of the road; and being in physical condition to drive, including avoiding alcohol impairment. The errors of novice drivers show the same general distribution as more experienced drivers, which is not surprising since the situations that lead to accidents, and the responses needed to prevent them, do not change with age or experience. What distinguishes the novices is the high initial accident rate, which affords the opportunity to achieve substantial benefit through DE.

Underlying many of the errors leading to accidents is failure to recognize a situation as hazardous. For example, a novice may see a car stopped at a cross street and fail to consider that the driver, looking the other way, might pull out, and thus be prepared to brake. The role of hazard recognition is difficult to infer from the information available from accident reports. Recognition of hazards differs from the other subjects in several ways:

1. It is almost entirely visual, calling for interpretations of the world as seen through windows and mirrors;
2. It is largely a perceptual skill, requiring immediate interpretation of visual stimuli as requiring a specific response, typically slowing or turning, rather than cognitive recall of some procedure; and
3. Since each hazard provides a unique visual image, recognition of a situation as hazardous requires some degree of generalization.

OBJECTIVES OF INSTRUCTION

Identifying the errors that lead to the high initial accident rate of novices helps chart an appropriate course of instruction for DE, one that focuses upon those involving the highest risk. The task becomes one of providing experiences in DE that lead to the same improvement as would occur on the road but without the same exposure to risk. One task will be to develop more realistic expectations for DE, which in earlier times was advanced as leading to a lifetime of responsible driving. The DeKalb project was criticized for a period of accident reduction lasting only 6 months. Actually very few interventions aimed at improving the driver performance show benefits beyond a year or so, after which everyday driving experience tends to bring everyone's performance up to the same level. But for novices, any educational effort that can reduce the extremely high risk of beginning driving will be of great benefit.

Safer driving can be sought through DE by the same means as other behavior changes—through development of needed knowledge, skills, and attitudes. Knowledge refers to stored information while skill implies ability acquired through practice. The meaning of the term attitude, however, ranges from a broad frame of mind to very specific beliefs. As has been noted, much of the early DE focused upon the first meaning, attempting to develop broad, responsible outlook on driving as a lifetime activity. More suitable objectives for DE courses are beliefs that support the specific various elements of driving that make up the course objectives. For example, novices must not only know they are expected to look both ways before crossing a main road but believe it is necessary to their own safety. In the absence of instruction, novices will acquire such attitudes through exposure to hazardous situations on the road, fortunately without ill effect—most of the time. Bringing them into DE requires means of safely exposing novices to situations capable of altering beliefs.

INSTRUCTIONAL CONTENT

When it comes to designing programs capable of achieving objectives consistent with the objectives described, DE instruction can be divided into three categories:

- Basic Procedures: vehicle control and maneuvering, laws and regulations;
- Safe Practices: search, attention, speed, space, signals, hazard recognition; and
- Advanced Skills: skid control, collision avoidance.

Basic procedures are a part of any driver instruction. Where teens are not required to complete DE, it may represent the sum total of instruction. However, for states mandating DE for licensing, instruction generally involves 30 h in the classroom and 6 h behind the wheel. Much of the latter is devoted simply to basic control of the vehicle. Once novices have had sufficient instruction to handle the vehicle, safe driving practices learned in the classroom can be brought onto the road. Instruction in skid control and collision avoidance requires facilities beyond what can be provided in most DE programs and is taught primarily by schools specializing in such instruction.

The challenge to DE is developing the knowledge and attitudes that underlie safe driving in ways that avoid the danger that arises when they are learned on the road. The next section will describe the means by which students can learn procedures and recognition of hazards through visual media capable of bringing elements of driving into the classroom and even the home. A program of hazard recognition addressing but a fraction of the safety threats encountered on the road has been shown to be effective in altering the driving of novices.

One development that opens the way for more effective DE is graduated licensing. The opportunity of dividing instruction into two phases allows basic driving procedures to be taught and mastered in the learner phase before attempting to superimpose the demands of various safe practices during the provisional license phase. While the merit of such an approach is obvious, it introduces some complications that must be addressed. As yet, only one state has made the attempt to do so.

Instructional Methods for Young Drivers

John F. Brock
General Dynamics

The most common method for teaching young men and women to drive is the traditional classroom lecture. The students sit at desks, the teacher stands at the front of the room lecturing, showing movies, and occasionally asking questions of the students. At some point, students are allowed to drive cars, many specifically fitted with right seat controls for the instructor. Typically, other students will be in the back seat of the vehicle.

This paper will discuss alternatives to traditional driver training techniques. Computer-based training (CBT) has a proven history of improving learning in other domains. Studies by the Institute of Defense Analysis suggest that CBT reduces instructional costs by about 1/3 and either reduces time of instruction by about 1/3 or increases the effectiveness of instruction by about 1/3. Current state-of-the-art driving simulators, a very specialized form of CBT, do not have equivalent data supporting their use. However, some forms of CBT that include part-task simulation have been shown to be effective. It is clear, however, that the future of effective young driver training will involve the application of advanced instructional technologies.

The trend of CBT is toward learner-centered design, which emphasizes the specific needs of the targeted learners, provides methods for the student to control the learning process and also measures student progress in terms of engagement (how does the student react to the instruction), effectiveness (how does the student perform at the completion of the instruction), and viability (what effect does the instruction have on the world in which the student is expected to perform). However, as was pointed out in the previous paper, without correct content the introduction of technology into the instructional process will have little or no effect. A famous aphorism in instructional technology is that the computer is to instruction what the refrigerated truck is to the lettuce in a grocery store.

Most young people are sophisticated users of technology. They also have high expectations of what computers should do (video games, high-quality graphics, and sound). But because they are users of technology, they should be receptive to well-designed and challenging CBT and part-task simulation programs. One can also envision using technology inside vehicles to both measure driver performance and to intervene before unsafe behavior leads to tragedy. Certainly, there is now an installed base of computer technology in schools and homes. With broadband access gaining predominance, the kind of graphics and realistic scenarios needed to simulate driving are more accessible. Not only young drivers, but parents and teachers can be helped by technology; electronic checklists and guidelines could be installed on handheld devices (including cell phones) and programs could provide illustrations of good and bad driving as well as provide guidance on parental involvement in the novice driver's learning process. The biggest barrier to all of this is a lack of dedicated funding.

Classroom instruction must be only a component of the training process; CBT programs, various levels of simulation, and in-car experiences will all be enhanced as the underlying technologies become ubiquitous and affordable. There is an obvious visual component to driving. Therefore, any technology-based instructional program must provide realistic and interactive visual experiences.

A fundamental question is how much of the actual driving experience can be taught and practiced through means other than in a vehicle. Traditionally, knowledge acquisition has been seen as appropriate for the classroom or technology-based instruction and the vehicle has been seen as the only place for practice. Recent computer-based programs and devices have attempted to change that formula; cameras, sensors and recorders, and ignition interlock devices are all examples of technology being used inside the car to provide feedback and corrective information to a driver—the traditional model for learning. Conversely, recent CD and DVD programs, as well as more expensive and complex simulators, have attempted to provide practice opportunities outside of the vehicle. Instructional technology must meet the individual needs of students; one advantage of technology is that it doesn't have to be a one-size-fits-all solution to student learning. By the same token, the optimum path to licensing and safe driving can vary by student. CBT, simulation, in-vehicle devices, and parental support can all be mixed and matched to optimize learning for every individual. The kind of analyses that would produce those capabilities is yet to be accomplished.

There is still a place for live instructors and more conventional media. A blending of simulation, classroom instructors, and individualized CBT is worth exploring. What is needed is an instructional design process that could match specific behavioral goals to instructional methods, order instructional content to maximize learning, identify appropriate tests to measure dynamic student progress, and ensure that when students finally get behind the wheel they are ready to learn how to drive safely.

In conclusion, it is clear that advanced instructional and assistive technologies can improve driver training. There is every reason to believe that CBT will become significantly more powerful, allowing improved visual simulation, a wider range of instructional content and activity options, and improved student performance measurement and recordkeeping capabilities. At the same time, technological innovations in vehicles (e.g., the Global Positioning Satellite system and Bluetooth) will provide better training and performance measurement in vehicle driving capabilities.

Student Competency Measures

LARRY LONERO
Northport Associates

Focusing the content of DE on critical skills and avoidance of the most common errors offers the opportunity to improve novice drivers' high initial accident rate. Key error categories involve visual search errors, attention errors, and speed choice errors with less frequent but important errors involving maintaining space, handling emergencies, basic car control, following rules of the road, and being in adequate physical condition. To be sure that appropriate skill objectives are being taught and mastered, valid, reliable measures of students' learning outcomes and competencies are needed. In addition to use as operational standards, such measures are also important as intermediate indicators for improving program development and evaluation, without diminishing the importance of improved safety as the overriding goal of DE.

CAPABILITIES AND BEST APPLICATIONS OF VARIOUS MEASURES

Written Tests

The principal function of written tests in DE is to reliably sample the knowledge that defines its objectives, and many knowledge tests are associated with curriculum materials. These tests are not typically standardized or validated by psychometric research. Research has been addressed to some other kinds of testing, such as inventories tapping psychological constructs (e.g., sensation seeking, hostility), which have shown moderate association with crashes. Testing for diagnostic assessment and direction to special needs education–training holds promise. It is less clear that psychological tests could be used as criterion measures for course completion, even in principle. Some relevant traits that might be tested for, such as maturity or responsibility, may be too situation specific and unstable to test reliably. Standardized tests with suitable norms and known psychometric properties should be developed and adopted.

Drive Tests

Tests of skill capabilities and measures of actual driving behavior can serve as evaluation criteria to ensure that skills objectives of DE are being met. Specialized and experimental drive tests have shown reliability and validity, but such tests seem to have had little operational implementation as course criterion measures. Passing government road tests is a traditional objective for DE, but these tests have little known reliability or validity, and there has been relatively little development work on them in recent years. Unobtrusive onboard monitoring of driving performance during or subsequent to training may also hold promise. Again, standardized tests with suitable norms and known psychometric properties are needed.

Simulation Measures

Computer-based training and testing (CBT&T) and simulation both hold promise for reliable and valid testing, but this promise has yet to be demonstrated empirically. Diagnostic competency testing and branching students into needed training can be integral in CBT&T, which can also permit self assessment and demonstration of individual limitations, but these potential benefits have not yet been realized in DE. We can differentiate approaches to simulation. First is closed-loop—scenes respond to driver as they respond to scenes. The second, open-loop simulation, presents fixed scenes to which drivers respond. The ability of video to present complex and dangerous traffic scenes allows driver decision making and hazard recognition to be tested inexpensively in classroom or by computer. It would seem to provide a means for both teaching and testing critical skills.

Other Testing and Measures

Teacher-based evaluation of in-class performances can be assessed and provide valuable information for developing student competencies. Focus groups and other qualitative approaches can have value in formative evaluation and program development, but they appear to be little used.

Setting Competency Standards

What Form Should They Take?

Standards should resemble learning objectives—what the student will be able to do in concrete behavioral terms. Course failure is unlikely to be acceptable to governments or parents, and failing students has negative business implications for DE programs. Other approaches might include incentives for passing—e.g., discount on cost of course. Test failure need not be final but could prompt more DE and/or remediation. Further incentives to good test performance might be possible—e.g., discount on minimum holding period for learner license.

Who Should Set Them?

Government-mandated standardized testing exists in other educational settings and might be possible in DE. However, this is a remote possibility until the low state of knowledge and lack of uniformity across and within jurisdictions is remedied. A range of largely voluntary standards from professional and trade associations is the more likely interim approach.

How Should They Be Set?

As much as possible, standards should be based on empirical research. They should address important factors in crashes. Suitable research is currently lacking and is much needed. Standards based on common sense, face validity, and expert opinion should be used only as temporary expedients and with great caution. The goal is a reliable, valid set of measures that operationalize driver competency.

Gearing Standards to Licensing Requirements

DE competency standards should be aimed higher and be more inclusive than minimal driver licensing standards. This could be permitted through greater amount of time available for testing in DE than in licensing. Standards could be required by licensing authorities and applied by DE programs as the U.S. oversight of DE is moving from departments of education to licensing agencies. Competency standards could also be used to govern progression through graduated driver licensing (GDL) stages. Individual jurisdictions will ultimately decide on standards, but experience has shown that excessive non-uniformity can result. There are probably not 50 or 60 different good ways to structure DE and GDL, so strong federal and association leadership is important.

NEEDS AND FUTURE DIRECTIONS

- Research and development (R&D) should be systematically supported to develop psychometrically reliable and appropriately validated standardized written tests, driving measures, and simulator protocols.
- For DE program operations, suitable measures should be developed for diagnostics and for deciding a student has achieved a passing standard of knowledge and skill.
- For evidence-based program development, measures of learning and performance outcomes need to be developed to form a part of the basic theory or logic model of the program. They should constitute intermediate objectives between program operations and the program's ultimate safety goals.
- Measures of performance capability and behavior should be validated against safety criteria. When this validation has achieved a modest level, intermediate measures of abilities and behavior should be used for evaluation criteria for improving and comparing programs.

Novice Driver Training Effectiveness Evaluation

RAYMOND C. PECK, SR.
RC Peck and Associates

Few would argue with the proposition that an effective driver training program should lead to a demonstrable reduction in the crash rates of novice drivers. In fact, the insertion of behind-the-wheel training into the high school curricula of many states in the 1950s and 1960s was largely based on claims that on-the-road driver training would produce substantial reductions in the crash rate of teenagers. An accumulation of evidence from a variety of quasi-experimental and experimental studies published over the past 40 years has failed to substantiate the initial optimism.

By far the most rigorous study to date is the NHTSA-funded experimentally controlled study conducted in DeKalb County, Georgia. As noted in a companion paper by McKnight, the DeKalb study reported evidence of a small short-term decline in crash rates among licensed drivers. Given subsequent evidence summarized by McKnight concerning the steepness of the learning curve (reduction in per-mile crash rate) over the first few months of driving, the fact that any training effect would be short-lived is not unexpected. Nevertheless, the statistically significant crash reduction observed in DeKalb is not without ambiguities. First, the random assignment was compromised by the significantly higher percentage of drivers in the trained groups who obtained their license shortly after training or control group assignment. The treatment group comparisons among licensed drivers were therefore subject to self-selection biases. A second source of ambiguity relates to comparisons between the two training groups. There was no evidence that those assigned to the intensive comprehensive training module (SPC) had lower crash rates than those assigned to the standard training module (PDL).

Even if one dismisses the suggestive evidence of crash reduction reported in the DeKalb study, one can still not reject the possibility that driver training has some impact on crash rates, particularly during the first 3 to 6 months of licensure. Aside from the well-known admonition that the null-hypotheses cannot be proven, there is a very real possibility that driver training has a small effect on crash rates, but one that is below the detectability threshold. The DeKalb study, for example, was designed to detect a 15% reduction in 1-year crash rates with a confidence (power) of .80. But what if driver training reduces crash rates by, say, only 5%? A reliable detection of this small but possibly cost-beneficial reduction would require sample sizes approaching 100,000 subjects. An experimentally controlled driver training study of this size has never been conducted and is probably unfeasible. Large quasi-experimental studies, though sometimes feasible, do not have sufficient sensitivity for detecting small effects in the likely presence of residual biases and confounding.

USE OF ALTERNATIVE EVALUATION CRITERIA

A number of investigators have proposed intermediate criteria or proxy variables for assessing the effectiveness of driver training. The development of valid intermediate criteria needs to begin with an understanding of the distinction between driving behavior and crash involvement. Traffic crashes are not really a behavioral measure but are rather the infrequent stochastic consequence of complex interactive behaviors and external events. Most often, unsafe driving behaviors do not result in crashes, and even the safest behaviors do not preclude a driver from being crash-involved. Theoretically, the most valid and precise measure of the effects of driver training would be provided by inconspicuous observations of each trainee's actual in-traffic driving over a period of time and over a wide variety of situations. Such a measure would actually trump crash frequency as the ultimate criterion because it represents the behavior that driver training directly impacts and it reflects how a driver actually drives when not being tested. Assuming that the observational ratings were reliable and valid indicators of crash risk, crash reduction could be inferred from evidence that those receiving training drove more "safely" than a randomly constituted comparison group. Although this type of observation has been used in a couple of published studies, practical and logistic considerations would appear to limit wide-scale use.

To date, most research on the use of intermediate criteria has been limited to written knowledge and road test scores. There is evidence that driver training (classroom and on the road) results in slightly higher scores on state Department of Motor Vehicle (DMV)-type knowledge tests and on DMV and experimental road tests. Although neither of these test modalities has ever been shown to correlate with crash rates, this does not necessarily mean that training-mediated increases in the knowledge and skill domains measured by the tests could not result in crash reduction.

Another candidate proxy measure that has been studied is traffic violation or citation frequency. This is not really an intermediate measure since it reflects incidents that occur parallel to, and sometimes simultaneous with, crash involvement. Traffic violation rates have several advantages over crash-involvement rates: they are more directly related to the actual behavior of a driver, they are known to be significantly correlated with crash involvement rates, they require much smaller samples for adequate statistical power; and they can be considered relevant outcome measures irrespective of their relationship to crash-involvement rates. However, the relationship between traffic violations and crash rates is not sufficiently strong to permit one to conclude with assurance that reductions in traffic violations will lead to reductions in traffic crashes. Nevertheless, it can be legitimately argued that reduction in traffic violations is an outcome that favors the occurrence of crash reduction.

FUTURE RESEARCH DIRECTIONS

If driver training is to have a measurable effect on crash rates, it should impact those intermediate or mediating factors that are most highly associated with crash causation and which potentially are modifiable through training. As noted in the companion paper by McKnight, search and scan strategies, critical cue perception, and hazard recognition are particularly critical to crash avoidance.

Part-task driving simulators and computerized or video-based tests offer promise in constructing driver training effectiveness outcome measures. Other advantages over crash rates would be the relative immediacy of the results and the much smaller sample sizes required for adequate statistical power. The disadvantage is that one could always question whether effects on these intermediate measures translate into crash reduction and, if so, into how much reduction. If one attempts to counter skepticism by validating the tests against crash rates, the very quandary cited above reemerges.

In addition to serving as potential outcome proxy measures, performance tests can also serve as assessment tools to gage the progress and competency level of students at various exit points in the training process. This use of testing is addressed in a companion paper by Lonero.

An obvious limitation of any road test, even one that is reliable and comprehensive, is the inability to include task demands that expose the test subject to crash risk, and there are also limitations in measurement precision. Despite these limitations, there is some evidence from DeKalb that a carefully developed road test can discriminate to a moderate degree between drivers who receive formal behind-the-wheel training and those who do not. For example, those receiving the enhanced training (SPC module) scored about 10% higher than controls on the University of Southern California (USC) Road Test. This road test, in turn, was based on the HUMRRO drive task analysis, which identified skills and behaviors considered to be critical to driving a vehicle safely, including hazard perception. Unfortunately, the DeKalb investigators did not evaluate whether training-mediated effects on the USC road test correlated with effects on crash rates.

Another limitation of a road test or any posttraining test, whether given by a researcher, DMV or driver educator, is that it measures a driver's knowledge and skill level under inherently artificial conditions. The fact that a driver possesses the requisite knowledge and skill may have little to do with how that driver actually drives when not being tested. Actual driving behavior beyond the brief interval immediately following training is probably more influenced by attitudinal and maturational factors affecting risk assessment and concepts of personal vulnerability. Although GDL programs attempt to minimize the impact of immaturity and inexperience, a key challenge facing driver educators is how to modify attitudinal factors affecting risky driving and how to evaluate the impact of attitude modification strategies on actual driving behavior. Outlined below is a strategy for developing improved skill tests followed by some very preliminary thoughts on measuring attitudes and actual on-the-road behavior:

1. Identify the most critical drive task elements by a review of prior drive task and critical incident crash studies.
2. Review the components of the USC road test and identify which elements can better be tapped by use of part-task simulators or computer- or video-based tests.
3. Develop operational definitions and test specifications. Based on what is presently known it is expected that the performance test will emphasize search and scan strategies, perception, risk awareness, situational awareness, and accident avoidance skill.
4. Develop and evaluate test prototypes.
5. Modify test based on results of No. 4 and verify reliability and other psychometric properties.
6. Conduct large scale field evaluation studies using randomized experimental designs to assess effects of driver training on test performance.

Assuming that the tests have adequate reliability and sensitivity, sample sizes of several hundred per group may be adequate for detecting moderate effects. In addition to allowing comparison between groups within different types or amounts of training, it would also be possible to perform a pre- versus posttraining comparison in which each trainee serves as his or her control.

A question that arises in evaluating the effectiveness of a driver training program relates to the nature of the control or comparison group. Inclusion of a pure no-treatment control group presents a number of logistical and feasibility problems and is not always essential. Drivers assigned to a non-training control condition will necessarily seek some type of training and practice in order to become licensed. It is therefore not possible to ever compare a formal driver training program with a group who has not been trained, at least informally.

In DeKalb, the control subjects were either trained by friends, parents, or a private driver training school. One of the objectives of the control group in DeKalb was to represent how novice drivers would have been trained and licensed had the formal driver training programs (SPC and PDL) not been offered. However, any generalization of this estimate beyond the sampled jurisdiction (DeKalb) is problematic and is potentially compromised by reactivity artifacts inherent in experiments where subjects know they are being treated differently.

An alternative to the use of a no-treatment control is to assign comparison group drivers to some form of minimal training and attempt to show that the experimental or enhanced training is superior. The PDL group in the DeKalb study provided this type of baseline.

It will be noted that the above plan does not include anything about attitude measurement. The emphasis on crash avoidance as an experience-mediated skill does not mean that attitudes and maturity are unimportant. In fact, the only test found to correlate with accident rates in the DeKalb study was the Mann Driving Attitude Inventory (MDAI). The MDAI or similar tests might be useful in measuring attitude shift following exposure to a driver educational course (whether one could infer crash reduction from evidence of attitudinal improvement is, of course, problematic and the hypothesis would have to be validated by an independent R&D effort). The approach outlined in the next section addresses attudinal factors without directly measuring them.

INCONSPICUOUS OBSERVATION

As noted earlier, any performance assessment made in a conventional testing context is inherently artificial because the driver knows he or she is being tested. The test therefore measures how a driver can drive but not necessarily how a given driver actually drives when not being observed. If driver training is to reduce crash rates, it must have some effect on actual driving behavior and not just the skill components measured under testing conditions. Two approaches to measuring driving behavior are summarized below:

- Behavioral observation: This involves following and videotaping drivers who have completed training or a road test and rating their performance on an array of dimensions, including search and scan behaviors and unsafe driving maneuvers. Limitations include the relatively brief time period of the observations and the driver’s becoming aware they were being observed. Some jurisdictions might also have reservations about participating if the observations are surreptitious. The psychometric properties of the ratings and large scale operational feasibility of the approach require additional research. The identification and inclusion of a comparison group may also present logistical problems.
- In-car computers: Technology exists for installing computers that measure a variety of parameters (mean speed, speed variance, acceleration and deceleration rates, brake presses, and lateral movement). This approach would provide a number of precise measures sampled over a substantial period of time. Among the limitations are the absence of direct measures of critical

behaviors, such as search and scan and confounding due to highway and traffic convictions. However, some of the measures would indirectly reflect erratic driving, poor anticipation, etc.

Research should be undertaken to evaluate the relative merits and operational feasibility of the above approaches before embarking on a large scale validation.

I believe it is fair to speculate that not everyone will agree with the premise of inferring crash reduction from demonstration of training effects on intermediate criteria. If this premise is rejected, one could argue that a demonstration of improved skill and "safer" driving per se is sufficient to support a prelicensure driver training requirement.

The Future of Driver Education

Daniel R. Mayhew
Traffic Injury Research Foundation

A stated objective of DE is to produce safer drivers, typically defined as drivers less likely to crash. This is certainly the expectation of policy makers, the media, and the public, and is not surprising given that DE has strong face validity as a safety measure. What is surprising to many is that recent reviews of the evaluation literature have confirmed what the research community has known for some time: traditional DE and training programs have yet to demonstrate consistent attainment of safety objectives.

Despite its disappointing safety record to date, the path ahead for DE holds promise in achieving important safety objectives for several reasons. Much has been learned from research and program development since 1998 when the midyear meeting of TRB's Operator Education and Regulation Committee focused on Driver Education at the Crossroads. We have a better understanding of young driver crashes and the critical factors that give rise to their elevated crash risk. As discussed in companion papers by Preusser and McKnight, the elevated crash risk of young drivers is highest over the first months and miles of driving, when novices are the most inexperienced and unskilled. Not surprisingly, many of the errors leading to crashes in this high-risk period relate to inexperience in driving, including for example visual search and attention errors. This suggests that to be effective as a safety measure, DE should be more clearly focused on addressing the extremely high crash risk facing beginners over their first months and miles of driving.

We also have a much better appreciation for the strengths and limitations of DE and how we can improve its safety potential based on the available research evidence. Even though DE has generally not proven to be an effective safety measure, it has a number of important strengths that should be built upon for the future. DE and behind-the-wheel training are efficient means for beginners to learn how to drive and develop their driving skills in a controlled and safe environment—under the supervision of an instructor. It also prepares novices to pass the road test, thereby increasing their mobility, which is highly valued in society for economic and other reasons.

However, there is an important tradeoff between mobility and safety. DE provides a means for teens to drive but this exposes them to crash risk. Accordingly, a number of limitations of DE need to be addressed to minimize this crash risk. There is a need to expand the relatively short duration of DE, which typically involves 30 h in class and 6 h in vehicle taken over a few months or weeks just before the road test. There is a need to focus on the development of the skills most critical to safe driving performance in situations where young drivers are at highest risk, rather than on a broad range of knowledge and skills in a relatively superficial manner. There is a need to adopt contemporary teaching methods and principles that are compatible with the individual needs of today's young drivers.

Potential improvements to overcome the limitations of DE have been discussed in more detail in the companion papers so are only mentioned briefly here. They include adopting a multiphased approach to DE that better integrates with GDL, a safety program with proven effectiveness. This would involve at least two stages of DE with one initially focusing on basic vehicle control skills, and the other, after some driving experience has been gained, focusing on higher-order safety knowledge and skills, such as recognizing and reacting to hazards. The content of DE could also be improved by focusing on those errors that contribute most to the initial high crash risk of novices, by motivating teens to drive safely as well as by providing insights to counteract overconfidence that might be the unanticipated byproduct of taking DE. Immaturity, risk taking, and peer pressure are factors in young driver crashes, and they are especially serious ones, but the extent to which DE can deal effectively with such "lifestyle" factors needs further consideration.

DE should also use the best teaching methods and learning principles—for example, CBT and driving simulation that provides a protective means for exposing teens to hazardous driving situations that contribute most to their elevated crash risk. On-road teaching techniques, such as commentary driving, and in-car technologies to train and measure young drivers should also be considered. It is important as well to match the learning experiences to the novices' needs and skill level, which speaks to better testing and diagnostic assessments, for example, by means of improved drive tests and CBT and testing.

There are several encouraging initiatives in the field of DE that illustrate growing interest in improving DE. The NTSB convened a 2-day public forum and produced a report that reviewed the current state of DE. It discussed the extent to which DE is used, its strengths and shortcomings, and what can be done to improve it. Recommendations included reviewing current programs and determining which instructional tools, training methods, and curricula are consistent with the best teaching methods to reduce crashes, and determining the optimum sequencing, in conjunction with GDL, for educating teenagers on safe driving skills.

There has also been a renewed interest in evaluating DE, and Peck, in a companion paper in this circular, addresses issues critical to evaluation, including the use of intermediate criteria or proxy measures. In this regard, a project under funding from the AAA Foundation for Traffic Safety and BMW involved a comprehensive review and consultation process to produce a report on Guidelines for Evaluating Driver Education Programs. This is intended to provide practical information for how to conduct different types and levels of evaluations.

A final example of the growing interest in improving DE and the importance of evaluating programs comes from Australia where there are plans to implement a compulsory national education scheme for probationary drivers by 2007. This multimillion-dollar project involves developing a model curriculum and evaluating it, using a scientifically sound and rigorous design.

Critical next steps for improving DE in the United States and Canada include a review of recent progress in the development of DE programs and the evaluation of promising new ones, especially those that have already adopted some of the features discussed in this circular. Research on young drivers and their crashes is also still needed to provide further guidance for improving DE so that it can achieve its safety objectives.

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

TRANSPORTATION RESEARCH BOARD
500 Fifth Street, NW
Washington, DC 20001