

THE RELATIONSHIP OF WRITTEN EXAMINATION PERFORMANCE  
TO SAFE DRIVING: A LITERATURE REVIEW WITH  
RECOMMENDED METHODS FOR DEVELOPING EXAMS

by

Alden L. Atkins  
Graduate Legal Assistant

(The opinions, findings, and conclusions expressed in this  
report are those of the author and not necessarily those of  
the sponsoring agencies.)

Virginia Highway & Transportation Research Council  
(A Cooperative Organization Sponsored Jointly by the Virginia  
Department of Highways & Transportation and  
the University of Virginia)

Charlottesville, Virginia

April 1984  
VHTRC 84-R41

SAFETY RESEARCH ADVISORY COMMITTEE

- W. E. DOUGLAS, Chairman, Director, Planning & Programs Development,  
Division of Motor Vehicles
- P. L. ASH, JR., Chief of Police, Staunton, Virginia
- V. M. BURGESS, Transportation Safety Administrator, Division of Motor  
Vehicles
- C. F. CLARK, Driver Services Administrator, Division of Motor Vehicles
- C. P. HEITZLER, JR., Program Manager, Department of Management Analysis  
and Systems Development
- B. G. JOHNSON, Supervisor, Driver Education, Department of Education
- C. S. JOHNSON, JR., Field Supervisor - West, Department of State Police
- R. F. MCCARTY, Safety Program Coordinator, FHWA
- W. F. MCCORMICK, Assistant District Engineer, VDH&T
- R. M. MCDONALD, Project Director, Transportation Safety Training Center,  
Virginia Commonwealth University
- S. D. MCHENRY, Director, Bureau of Emergency Medical Services,  
Department of Health
- F. F. SMALL, Highway Engineering Program Supervisor, VDH&T
- J. A. SPENCER, Assistant Attorney General, Office of the Attorney  
General
- C. B. STOKE, Research Scientist, VH&TRC
- E. W. TIMMONS, Director of Public Affairs, Tidewater AAA of Virginia,  
Norfolk, Virginia

## ABSTRACT

This report reviews the literature concerning written driver license examinations. The research literature shows that current written examinations are poor predictors of unsafe drivers. Although some studies demonstrate significant relationships between one's written examination score and accidents, these relationships are significant only for drivers with certain combinations of sex, age, and level of education. Even for those classes of persons where a significant relationship was found, failing examinations is very over-inclusive. Thus, many safe drivers would have to be failed to screen out one unsafe driver. The conclusions, however, are based upon tests which are currently used and which have been widely criticized as not clearly testing knowledge or as not being statistically reliable. Another reason to administer written examinations is to mend information deficiencies. This can be particularly effective if drivers can be classified into groups with identified information deficiencies. This report also reviews classifications with identified information deficiencies, and it suggests how further research can be conducted with properly developed examinations.



THE RELATIONSHIP OF WRITTEN EXAMINATION PERFORMANCE  
TO SAFE DRIVING: A LITERATURE REVIEW WITH  
RECOMMENDED METHODS FOR DEVELOPING EXAMS

by

Alden L. Atkins  
Graduate Legal Assistant

PURPOSE AND SCOPE

The Virginia Division of Motor Vehicles (DMV) proposes to revise the written examination given applicants for driver licenses. In anticipation of the revision, the DMV has commissioned the Virginia Highway and Transportation Research Council to study several aspects of driver testing and licensing, the first phase of which is a review of the relevant literature. This report reviews the literature, focusing on current knowledge concerning the purposes of written examinations, their ability to predict or alter unsafe driving, and methods of developing examinations. It concludes with a description of how examinations could be developed for further research.

DISCUSSION

The Highway Safety Act of 1966<sup>(1)</sup> requires all states to follow highway safety standards promulgated by the Secretary of Transportation. These standards require the applicant for a driver's license to demonstrate a comprehensive knowledge of traffic signs and a "knowledge of laws relating to traffic (rules of the road) safe driving procedures, vehicle and highway safety features, emergency situations that arise in the operation of an automobile, and other driver responsibilities."<sup>(2)</sup> Thus, drivers must be tested for knowledge when they first apply for a license and every four years thereafter. Accordingly, Virginia requires all applicants, except those seeking a renewal who have had fewer than two citations in the preceding four years, to pass a written examination.<sup>(3)</sup>

The purpose of written examinations is not particularly clear. The regulations promulgated under the Highway Safety Act say that their purpose is to ensure that drivers are "mentally and physically qualified."<sup>(4)</sup> Similarly, the Virginia Code requires the DMV to test a driver's "physical and mental qualifications and his ability to operate a motor vehicle in such manner as not to jeopardize the safety of persons or property."<sup>(5)</sup> This implies that the purpose of written

examinations is to identify unsafe drivers and prohibit them from driving.

Researchers have suggested other purposes for written examinations. One is that the license fee raises revenue for the state. Waller(6) points out, however, that "most administrators would neither wish to defend them on that basis or find it easy to do so, since there are far more efficient means for collecting revenue."(7) Another purpose is that they encourage drivers to learn the rules of the road and any other information included in the examination. Coppin(8) goes a step further by suggesting the use of examinations to diagnose driving problems which the state can correct by directing the driver to proper remedial support. Waller argues that the diagnosis should not be limited to driving deficiencies, but should extend to a brief physical examination of medical problems as well. The sections which follow consider the research literature which supports the different purposes for written examinations.

#### Written Examinations as Predictors of Unsafe Driving Behavior

Written examinations are often justified as a means of identifying unsafe drivers. The underlying assumption is that knowledge of the information tested lessens a person's chances of having an accident. The difficulty with this assumption is that it is not clear that greater knowledge actually leads to fewer accidents.

Malfetti and Fine(9) studied the attributes of the safe driver. They administered a battery of tests to six professional drivers, each of whom had not had a preventable accident in the preceding 20 years. Included in the tests was a written examination lasting 4 hours and designed to test knowledge of the rules of the road and of general highway safety. Significantly, five of the six exceptional drivers scored "average" to "poor" and only one scored "good" on the written examination. Although these results cannot be generalized because of the small number of subjects, exceptionally safe drivers do not necessarily have better than average knowledge of driving principles. Instead, the study observed, "during the performance tests the subjects seemed to use what knowledge they possessed to a greater extent than do most drivers."(10) The study concluded that behavioral characteristics, such as nonaggression and being sensitive to others' opinions, were more important to safe driving than knowledge.

This is not to say that knowledge is entirely unrelated to safe driving. A few studies have found a weak, though statistically significant, relationship between driving and scores made on written examinations. Still other studies have not found this relationship to be statistically significant. The following is a discussion of these

studies and whether written examinations can predict driving performance.

### The Limited Predictive Ability of Written Examinations

Most of the research literature concludes that written examinations cannot accurately predict driving behavior. This is true with respect to being involved in accidents as well as being issued citations. Further, two studies have proposed that licensing procedures be streamlined to save money because they conclude that this will not affect highway safety.

Stoke(11) studied the relationship of performance on a written examination with subsequent accident and citation history. Under a waiver of the nationwide licensing standards by the U. S. Department of Transportation, applicants were divided into four groups: (1) a control group who did not take a written examination, (2) a group which received only the driver's manual, (3) a group that received the manual and an exam to be taken at home, and (4) a group that received the manual and took the exam at the DMV examining station. Although Stoke found a few statistically significant relationships among the 241 comparisons made, he concluded that examinations could not predict accidents or citations because the few significant relationships revealed no discernible pattern. The study proposed that written reexaminations be eliminated because they do not produce cost savings which offset the costs of administering them. This study showed, first, that examinations on the rules of the road cannot adequately screen drivers, and second, that even if they do create an incentive to learn the material in the driver's manual, the knowledge has no effect on subsequent driving. Assuming that knowledge is of some importance, either the written examinations do not create proper incentives or alternatively the examinations and manuals do not cover the proper material.

Most written examinations cover only the rules of the road, however, Carpenter(12) included questions about safe driving as well. He then evaluated the ability of the expanded exam to predict driving behavior for new drivers, drivers with out-of-state licenses applying for California licenses, and "intermediate" (fair and poor) drivers. He found that passing the expanded exam had no effect on out-of-state drivers or on "intermediate" drivers. Surprisingly, new drivers who passed the expanded exam actually had more subsequent accidents than those who passed the rules of the road exam. Although whether one passed or failed generally had no effect on accidents, the studies did find a weak correlation between the number of questions missed and subsequent driving. Even so, Carpenter concluded that the \$150,000 cost to implement the expanded exam was not justified because any resulting accident reductions were conjectural. This study showed that testing

knowledge of safe driving principles in addition to the rules of the road did not add to the predictive ability of the exam.

In a similar vein, Conley and Smiley(13) studied the predictive ability of examinations with questions designed to test knowledge relevant to eleven categories of traffic violations. Although the study found that performance on certain test forms was correlated to different subsequent citations, it concluded that as a general rule knowledge cannot predict the number of citations one will later receive. Assuming that an unsafe driver receives more citations than a safe driver, this study indicates that a written examination cannot adequately screen unsafe drivers.

Freeburg and Creech(14) studied North Carolina's written examination, and they compared performance on the examination to past rather than future driving performance. They found that test score alone did not accurately predict accidents or citations. Only when background information such as age, formal education, years of driving, and miles driven per week was added to the correlation could the examination score identify unsafe drivers, and even then it was successful only with men. In a similar study, McKnight and Edwards(15) designed special examinations for drivers with certain characteristics, specifically new applicants, renewal applicants, and older applicants. Their study found a significant relationship between accidents and written test score for new drivers, for renewals with prior accidents or citations, and for some older drivers. They did not find, however, that the examinations could identify unsafe drivers from all backgrounds. Thus, if the purpose of the examination is to identify unsafe drivers, the test score would have to be combined with background information to determine who had "failed" the examination. Such a discriminatory practice could not be politically justified.

It is important to note that McKnight and Edwards's study is aimed not at identifying unsafe drivers but at diagnosing driver problems which certain types of drivers have. Drivers' problems can then be combatted by providing appropriate information. Thus, their study has shown that although generalized knowledge may not have a safety effect, creating an incentive to learn information commonly deficient in certain types of drivers can have a safety effect. The study does not show, however, that examinations can be used to identify unsafe drivers.

Only Dreyer(16) concludes that written examinations can predict accidents. His study developed exams which tested only knowledge of the rules of the road and of signs; it did not include questions covering highway safety. The study concluded that those with better written test scores tended to have fewer subsequent accidents and convictions; however, these correlations were low.

These studies together show that knowledge of the rules of the road and of highway safety information may be important but not absolutely necessary to safe driving. Instead, knowledge is probably only one of many attributes of a good driver. One who passes a written examination may be an unsafe driver because he may lack one or more other important characteristics. Although knowledge is important, it is more important for drivers with certain characteristics. Important variables are sex, age, education, and prior driving history. Thus, if written examinations can predict future driving behavior at all, their ability to do so is limited to drivers with these characteristics.

#### Practical Reasons Why Written Examinations Cannot be Used To Identify Unsafe Drivers

As noted above, although performance on written examinations may be related to safe driving, the relationship is at best a weak one. Assuming that a weak relationship exists, there are three significant problems with using written examinations to keep unsafe drivers off the road. First, the resulting change in the number of accidents would be very small. Second, the public would not tolerate a significant increase in the number of persons denied a license. Third, the courts may not find that such a weak relationship is a rational basis for denying one the right to drive.

There are several studies which indicate that removing unsafe drivers from the highways would have little impact on the number of accidents each year. Waller discusses a series of studies which have concluded that removing all drivers who have been involved in an accident, people who are traditionally considered unsafe, would have little effect on the total number of accidents. (17)

In other words, most people involved in accidents during a specified period are not people who are unsafe as measured by prior accident history. Similar results will occur any time the criteria used to identify unsafe drivers either (1) are not strongly correlated to unsafe driving, or (2) make up only a few of the many characteristics which make one an unsafe driver. As shown above, both of these problems apply to performance on a written examination. McRae (18) used assumed figures consistent with the findings of his study, and he pointed out that 35% of all future good drivers would have to be failed in order to fail 50% of future unsafe drivers.

This problem is caused by the fact that there are two types of errors when using written examinations to identify unsafe drivers. First, the exam will be over-inclusive by incorrectly predicting that some drivers will be unsafe when they would actually drive safely; and second, the exam will be under-inclusive because some drivers who pass

the exam will not drive safely. If the purpose of the examination is to screen unsafe drivers, the number of people incorrectly identified as safe must be minimized. Because this can be done only by making the exam even more over-inclusive, a large number of people who would be safe drivers would have to be failed. McRae concludes, under the assumptions of his example, that 40 future good drivers would have to fail for every 1 future bad driver who fails. Given the extent to which our society depends on the automobile, the public would quickly react to such a high failure rate and complain to political representatives. Either because of the political pressure or a change in the state laws, the DMV would have to change its high failure policy.

The political ramifications would be even greater if the DMV were to vary the passing score with a driver's characteristics in light of the research which shows that examinations are better able to predict subsequent driving behavior of certain types of drivers. This would quickly lead to accusations of discrimination, and again the DMV would be pressured to change its policy.

Further, Waller argues that the courts are reaching the point where they might order such a policy to be changed. Although the states view driving as a privilege, the courts are beginning to view it as a right. If driving is a right, then the safeguards of the due process clause of the Fourteenth Amendment apply to it. In Bell v. Burson, (19), the United States Supreme Court reached just such a conclusion when it required Georgia to give notice and a hearing before suspending the defendant's license. The Court went on to say,

[o]nce licenses are issued, ... their continued possession may become essential in the pursuit of a livelihood. Suspension of issued licenses thus involve state action that adjudicates important interests of the licensee. In such cases the licenses are not to be taken away without that procedural due process required by the Fourteenth Amendment .... [citations] This is but an application of the general proposition that relevant constitutional restraints limit state power to terminate an entitlement whether the entitlement is denominated as a 'right' or a 'privilege'. (20)

Similarly, this reasoning could be applied to rejecting applicants as unsafe because they failed written examinations when research indicates that there is little or no rational basis for believing that safety and knowledge are related.

## Written Examinations as an Incentive to Learn the Rules of the Road

As noted earlier, several researchers have agreed that the purpose of written examinations is to give applicants an incentive to learn the materials in the driver handbook. This section will discuss first whether this purpose can be justified and then what information should be tested.

### Justifying the Incentive to Learn

If the purpose of written examinations is to create an incentive to learn, there must be a benefit to knowing this information. Underlying this rationale for requiring written examinations is the unarticulated assumption that having this knowledge makes one a better driver. The research does not clearly indicate whether this assumption is justified.

Most of the studies which indicate that having such knowledge does not have a safety benefit are ones which show that performance on a written exam does not predict unsafe driving. It can be argued that these studies show that knowledge is not important because, assuming that the tests reliably measure knowledge, those who fail the examination drive as well as those who pass. These studies can be distinguished, however, because this rationale (creating an incentive to learn) is not concerned with one's performance on the examination. Instead, this rationale is interested only in failing enough people to create a real incentive for applicants to study the driving manual. Arguably, even those who fail the examination have studied the driving manual, and thus one would expect their subsequent accident history to be no different from that of those who pass.

Instead, the more relevant studies compare the subsequent driving behavior of persons who have taken a written examination with the behavior of those who have not. Helander<sup>(21)</sup> conducted a study involving drivers who had had a serious accident problem in the preceding one-year period. California law requires such persons to be reexamined to determine if the driver's operator permit should be revoked. Helander compared the subsequent accident and citation histories of drivers who had been required to be reexamined and those of a control group of drivers who had not been reexamined. He found that subsequent accidents fell a statistically significant 21% when the examination was required. He also found, however, that subsequent citations rose a small, but not statistically significant, amount when the examination was required. Although this study measured the effects of a program that informed the unsafe driver that the state was concerned with his driving, it also can be taken as a measure of the incentive effects of learning the material to pass a written examination. Stoke, however,

compared the subsequent accidents and citations of persons who were asked to take an examination with those who were not, and he found no significantly different subsequent driving behavior. It is important to note, however, that applicants in Stoke's study did not have to pass the written examination to be relicensed. Although many applicants may have thought that they had to pass the examination, the media publicized the fact that applicants did not have to pass the test before the study began. As a result, many applicants refused to take any written examination. In practice, however, applicants who fail the examination may not have their license renewed. Thus, the incentives for applicants to learn the material in the driving manual in Stoke's study may have been less than the incentives both in Helander's study and in practice, where applicants may lose the right to drive if they fail. As a result, it is not clear whether written examinations have a safety effect by creating an incentive to learn the material to be tested. Significantly, however, this justification for the examination does not have the problems with being over- or under-exclusive that using examinations to screen unsafe drivers has.

#### Information Which Should Be Included

Traditionally, the written examination in most states has included questions about only two subjects: traffic signs, signals and markings, and the rules of the road. The rules of the road are the state's traffic laws. Because the examinations tend to cover only information included in the driver manuals, the manuals similarly tend to focus on these two subjects. A growing body of literature suggests, however, that the manuals and examinations should include safety issues as well. The inclusion of safety information would be consistent with viewing written examinations as creating an incentive to learn, because drivers would learn information relevant to the safe operation of a motor vehicle. This section shall discuss what types of information should be included.

Current examinations have been criticized because they do not clearly test anything. Freeburg and Creech<sup>(22)</sup> criticize North Carolina's written examination because

[t]here was, unfortunately, no clear meaning derivable from the item content that could be assigned to any of the groupings, for any of the factors, on any of the test forms. This result is not entirely surprising if (as is likely) the test items were selected solely to provide maximum coverage of the material in the Driver's Handbook and with no a priori intention of selecting sets of items that cover particular subject matter (e.g., use of signals, maintenance of distance, emergency procedures, etc.).<sup>(23)</sup>

This problem can be eliminated by planning the subject matters deemed important in advance and revising the handbook accordingly. It should be noted, however, that Carpenter found that adding safety questions to the standard examination had no subsequent behavioral impact.

Reel(24) revised Michigan's written examination with the view that its purpose was to create an incentive to learn, and he expanded the subjects covered. He established six categories of information: (1) rules of the road, (2) safe driving procedures, (3) accident information, (4) licensing regulations, (5) alcohol and implied consent, and (6) vehicle equipment. Coppin suggests that an examination should cover those subjects as well as emergency procedures, the differences between urban and rural driving, pedestrian safety, cyclist safety, and the interaction of the vehicle with highway features and community environment. In this way, applicants can be given an incentive to learn not only the rules of the road but how to drive safely in varying situations as well. Similarly, Conley and Smiley(25) studied the predictive ability of an examination which had been revised to cover subjects deemed relevant to the violations being studied. They determined what knowledge might be relevant to predicting whether a driver would violate a particular law, and they designed the examination accordingly. If the DMV decides to study the predictive ability of its examinations, it should use a similar approach.

McKnight and Green(26) took this approach a step further by identifying information relevant for applicants with different characteristics. They divided the applicants into seven target groups -- new drivers, youthful drivers, renewal applicants, older drivers, traffic violators, accident repeaters, and drinking drivers -- and designed written examinations and driving manuals according to the perceived information needs of each group. They determined that new and youthful drivers could be combined into one group because they both required comprehensive information. This information related to licensing procedures, observation, means of communicating with other drivers, speed control, inter-vehicle separation, gap judgement, handling emergencies, physical and psychological factors, and vehicle maintenance. Experienced drivers renewing their licenses have acquired most of their knowledge from experience, and although the study found that their information strengths and deficiencies were similar to those of new drivers, the study deleted from the examinations information generally known by experienced drivers. This deleted information related to traffic signs and control signals, right-of-way laws, turn signals, and estimating gaps. Older drivers were the only group that McKnight and Green found had special information needs. Older drivers were informed of the health dangers often experienced by the elderly including fatigue, reduced eyesight and hearing confusion, and driving while under medication. They were also informed of specific unsafe driving habits

that older persons often develop; these were driving too slowly, failing to look to the rear, not understanding new traffic signs and laws, and failing to use alternative modes of transportation. Violators, they found, had characteristics very similar to those of non-violators, and as a result the violators' only information deficiency was found to be not the traffic laws themselves but the reason for the laws. When informed of the rationales for the laws, violators would understand the safety consequences of violating the law and commit such actions less often. McKnight and Green found that most drivers involved in accidents view themselves as victims rather than perpetrators, and thus accident repeaters required defensive driving information aimed at helping them avoid accidents happening to them. Finally, drinking drivers were found to have a sufficient knowledge of the traffic laws, but they did require special information to help them reduce their consumption of alcohol.

Many of these studies involve the design, pretest, and experimental use of one or more test forms. Unfortunately, however, only one of the reports on the studies included the test items that were developed. Carpenter is the only exception, and even he included only one of many test forms in each study.(27) The reason that the items are not generally included is probably that given by Freeberg and Creech -- to retain test security. As a result, the DMV may have to develop and pretest its own item pool if the researchers cannot be persuaded to release their test forms. Appendix A contains a list of the reports on these studies with their author's addresses.

As a result, the Virginia DMV should at least begin considering what types of information should be included in its revised examinations. If it adopts the position that the purpose of the examination is to create an incentive to learn, it should expand the coverage of the examinations and manuals to include more than merely traffic laws and signs. The Virginia Code allows this expansion because it prohibits testing only for information that does not pertain to the safe use of an automobile.(28) Reel's six categories of information discussed earlier (rules of the road, safe driving procedures, accident information, licensing regulations, alcohol and implied consent, and vehicle equipment) provide an appropriate place to begin considering what information should be included. If a more ambitious revision is deemed desirable, the DMV could adopt McKnight and Green's approach and target the subjects tested to drivers with different characteristics. If it does so, the DMV may want to establish its own classifications. In either event, the drivers manual should be revised concurrently with the examination in order to provide a source of the information.

## Pretesting the Examination

Any revised examinations which the DMV develops should be pretested before being administered to the general public. This is particularly true if the revision will be done with a view to subsequent research. Otherwise, the results of the examination may be inconsistent, thereby failing some examinees unjustly and making research conclusions impossible.

First, the average applicant should be able to understand the examination questions. This involves a two-part process of first testing the reading level that the examination requires and second asking readers if questions are ambiguous. The process is an iterative one which requires any revisions for clarity to be tested for reading level and vice versa. The appropriate reading level demands a policy decision by the DMV. Conley and Huffman chose a ninth grade reading level because 82% of the Illinois (the state where they conducted their study) population had completed eighth grade. Flesch(29) provides a method to test the reading level that a written work demands. First, random samples 100 words long should be selected, or if a piece is short, the entire piece should be used. Second, the average number of words per sentence should be calculated and rounded off to the nearest whole number. Third, the average word length should be calculated by dividing the number of syllables by the number of words. The reading ease score is then determined by the following formula:

$$\begin{aligned} & \text{Average Sentence Length} \times 1.015 \\ & + \text{Average Word Length} \quad \times 0.846 \\ & - 206.835 \\ & = \text{Reading Ease Score} \end{aligned}$$

The score means as follows: 100-90 very easy (5th grade); 90-80 easy (6th grade); 80-70 fairly easy (7th grade); 70-60 standard (8th and 9th grade); 60-50 fairly difficult (10th to 12th grade); 50-30 difficult (college) 30-0 very difficult (college graduate). Conley and Huffman used exactly this process to ensure that the exams they developed were of the appropriate level of difficulty.

Having tested an exam's reading level, the next step is to prevent ambiguity. McKnight and Green established the following rules to ensure that their test would not be ambiguous: (1) the questions and answers were stated briefly and clearly, (2) negative questions were avoided, (3) alternative answers used similar terminology and were of similar

length, (4) safety buzzwords were avoided, (5) questions were expressed in terms of driving application as often as possible, (6) keywords were emphasized, and (7) ambiguous words were avoided. Further, they formulated the alternative responses to each question using the following rules: (1) the alternatives all dealt with the same item of information, (2) only one response was correct instead of requiring the examinee to determine the best answer, (3) none of the alternatives were mutually exclusive, and (4) the answers "all of the above" and "none of the above" were avoided. Ambiguity can be further tested by asking pretest examinees to identify unclear questions or answers.

Next, the questions should be neither too easy nor too difficult. This can be tested only by administering the questions to a random sample of examinees. Conley and Huffman adopted a standard that no more than 70% nor less than 40% of all examinees should answer the question correctly. Although they found they could not rigidly apply this standard, they did reject items which deviated from it significantly. Further, they rejected or revised "foils", incorrect possible answers, which less than 2% of the examinees chose.

Once questions have been properly formulated, they must be assembled into a series of test forms. The first concern with the examination forms is that they be reliable. In effect, this means that no test form should be either significantly more or less difficult than the other test forms. A simple method of testing reliability is to randomly assign examinees to take different test forms and then compare the failure rates between forms for statistically significant differences. A more rigorous method, used by Dreyer<sup>(30)</sup> and Freeburg and Creech, is to use coefficient alpha. Ultimately, these tests of reliability show whether the examination results are reproducible. Reliability is a necessary, but not sufficient, condition for validity. Validity is the degree to which the examination measures what it hopes to measure. Dreyer points out that "[t]he external criterion that the licensing process must ultimately influence or relate to, is the applicant's driving record."<sup>(31)</sup> As discussed earlier, this correlation is at best weak, which means that it will probably be impossible to formulate a truly valid examination. If, on the other hand, the examination's purpose is solely to measure knowledge rather than to screen unsafe drivers, comparing the test form score to the score on a more comprehensive examination could establish validity. This would be consistent with viewing the purpose of written examination as providing an incentive to learn; however, the pretesting would be more costly because each examinee would have to take two examinations.

Assembling the items into different test forms requires careful planning. McKnight and Green first made sure that all items had been included in the driver's manual. Then, they selected the items based on three criteria: comprehensiveness, level of difficulty, and internal

consistency. To make the examinations comprehensive, items covering all of the desired categories of information must be included. McKnight and Green also varied the level of difficulty of test items. Although they did not exclude particularly easy or difficult questions as was recommended above, their approach is useful in that it suggests varying the difficulty of the questions. This can be done simply by studying the failure rate on each item. Finally, they made the forms internally consistent by not including items on the basis of their correlation with other items. Often test forms have several items in common, and although close similarity between forms makes for reliability, it also defeats the purpose of multiple forms.

Finally, test forms must be constructed with a concern for length. Freeberg and Creech noted that North Carolina had fewer questions on renewal test forms than on those for new applicants, and as a result the shorter examinations were less reliable. For this reason, they recommended that the test forms be lengthened to 30 or more questions. This would be particularly true if the examination results are to be used later to study the correlation of examination score to subsequent driving history.

#### Examination Methods

The examinations, once formulated, can be administered in a number of ways and at a number of times. This section will consider first when examinations should be required, second the use of a home test, and third testing equipment.

Federal standards require that drivers be retested at least every four years; however, such frequent retesting may not be necessary. Kelsey and Janke<sup>(32)</sup> studied the effects of not requiring retesting after four years. They found that a single four-year waiver had no effect on a driver's subsequent driving record. Similarly, they found no basis for California's requirement that drivers over age 70 be retested every two years. Thus, they recommended that the four-year and two-year standards be doubled and that the effects of longer waivers be studied. It should be noted, however, that their study was limited to drivers without accidents or convictions during the preceding four years.

Similar research has studied the use of waivers as a "reward" for good driving. This is particularly important here, because Virginia allows the DMV to waive the written examination for applicants with fewer than two citations in four years.<sup>(33)</sup> Kelsey and Janke discuss the literature on this subject, and they note a study by Harano and Hubert which found that drivers given a one-year extension as a reward for safe driving had more accidents during the extension year.<sup>(34)</sup>

However, they also studied an incentive program which notified unsafe drivers that they would be given an extension if they drove without incident for a year. The incentive program significantly reduced accidents during the subsequent year. Further, Kelsey and Janke's own study shows that not testing drivers without a conviction for four years does not affect their continued safe driving. The only difference between these studies appears to be the fact that Harano and Hubert's study gave drivers an unexpected reward instead of giving drivers a safe driving goal in advance. The reward must focus on future behavior to create an incentive rather than solely on past behavior. As a result, an extension would be desirable for safe drivers because administrative costs could be reduced; however, if the extension is given as a reward it should not be given unexpectedly. Instead, applicants should be told in advance that they can earn an extension through safe driving.

Recent studies have considered home testing as an alternative to requiring the applicant to take the examination at the DMV. Stoke studied home testing, and he found that the subsequent driving records of those who took the examination at home were indistinguishable from the records of those who took the examination at the DMV. The only drivers who were significantly affected were those who refused to take any written examination at all. Home testing, however, would not allow the licensing process to screen drivers for health problems. As suggested in one article,<sup>(25)</sup> the application process affords an excellent opportunity to test more of the applicant's physical characteristics than merely vision. In particular, applicants can be tested for hypertension and diabetes, each of which can lead to an unexpected driving emergency. Further, such a program could diagnose applicants who were previously unaware of their disease, thus creating a general benefit to the health of the community.

Alternative means of testing have also been considered. Many states offer oral examinations for illiterate applicants. McKnight and Green estimate that between 10% and 20% of all applicants request and are administered oral examinations. They studied using different means to present and to test information using the following modes: written; graphic, where the standard written question was accompanied by a diagram; static audiovisual, with slides and a taped oral presentation; and dynamic audiovisual, or movies. Although they found that applicants retained more information when the dynamic audiovisual system was used, the costs of such a system would be prohibitive. Reel similarly discusses the use of ITEM (individual test and examination machines). ITEM is an interactive machine where each question appears on a screen, the applicant answers it by pressing the appropriate button, and the unit both records the response and gives the applicant the correct answer. With the recent rise in the use of personal computers, they could be used even more effectively than ITEM in interactive testing. Further, if the examination's purpose is to provide an incentive to learn, such

an interactive system has the advantage of making the examination itself a learning experience because it provides immediate feedback. Again, however, the costs of such a program would be significant.

Thus, the examination can be administered at a number of times and in a number of ways. The evidence indicates that applicants do not have to be retested every four years; however, if the retesting standard is retained, safety benefits can be created by informing drivers in advance that the retest will be waived if they drive safely. The evidence also indicates that home testing has no safety effects; however, it would not allow the DMV to expand the physical examination to characteristics beyond eyesight. Finally, there are alternative testing modes available, and although they lead to greater information retention, their costs would be very high.

### CONCLUSIONS

In revising its written examinations, the DMV has the opportunity to clearly define the purpose of the exam, to draft the exams accordingly, and to conduct research into the relationship of knowledge and driving behavior. Most of the research conducted to date has not shown that written examinations can accurately predict driving behavior. At best, written examinations may be able to predict driving only when the driver's characteristics, such as sex, education, and weekly miles driven, are considered as well. Even so, the DMV would have to increase the failure rate dramatically to adequately screen unsafe drivers. This would not detect all, or perhaps even most, unsafe drivers; however, it would have to fail many safe drivers. Because driving is widely viewed as a right, the DMV would quickly be pressured to eliminate the high failure rate by the general public, the legislature, and perhaps the courts. Thus, according to current research, written examinations cannot be justified as a means of screening unsafe drivers. This is not to say, however, that more refined research is not necessary.

The justification for written examinations is twofold: (1) to amend information deficiencies, and (2) to identify applicants with various deficiencies and direct them to remedial help. This can be done for the population as a whole, or drivers can be classified by personal characteristics. Although these classifications could be very particular, research indicates that they do not need to be any more specific than new drivers, renewal applicants, and older drivers. Having identified deficiencies, the DMV should draw up manuals and examination questions accordingly. The questions should be tested for the required reading level, ambiguity, and excessive or insufficient difficulty. Examination forms are assembled from these questions, and each form should not be too short, should cover all the important subjects, and

should include questions of varying difficulty. The test forms should then be tested for reliability to ensure that they are of equal difficulty.

In preparing the written examinations, the DMV also has an opportunity to further research into the relationship of performance on a written examination and driving behavior. Based on the results of prior research, the DMV can revise its examinations with a view to more refined research. The most significant weakness of the prior research arises from the fact that accidents and citations are rare events. They are, however, the best means available for objectively measuring unsafe driving. Because accidents and citations are rare events, a study must be of proportions that will ensure that it will include enough of them to provide statistically meaningful results. Thus, the weakness of most studies relates first to the sample size and second to the length of the period for which accidents and citations are recorded. Although a few studies have used a relatively small population, their more pervasive weakness was the length of the period. Most studies record accidents over a period of two years or less, and none of those reviewed included data for more than a four-year period. Thus, the DMV has the opportunity to overcome the problem that accidents and citations are rare events by studying a large sample over four or more years.

A related opportunity for research is to compare pre- and post-examination accident records. Most of the research focuses only on pre- or posttest accidents, and the DMV could meaningfully contribute to the research by measuring accidents and citations over a significant period before and after the examination. The studies reviewed also tended to selectively choose examination stations rather than using all examination stations statewide. By selectively choosing stations, these studies tended to introduce a bias between urban and rural applicants in the study population. Finally, the prior research studies rarely measured exposure. In particular, they have assumed that persons who drive more have more accidents and citations because they are exposed to the risk more often. These people may, however, be safer per mile driven because they have more driving experience. The DMV could remedy this weakness by finding a measure of exposure and including it in the research. One way of measuring exposure, for example, is to ask applicants to estimate how many miles they drive each week.

In light of these weaknesses in prior research, the DMV could research one or more of the following areas:

1. To determine whether expanding the examination has a safety benefit. This can be done by randomly assigning applicants to take a standard rules of the road examination (and use the accompanying manual) or an expanded exam (and revised manual). (36)

2. To determine whether examinations have an incentive effect on learning and whether this has a safety benefit. This can be done by waiving the examination as a renewal requirement for randomly selected applicants and comparing the groups for differences in prior and subsequent driving records. (37)
3. To determine the relationships of certain areas of knowledge and certain types of accidents and citations. This involves: (1) grouping the questions by subject, and scoring each applicant for each subject, (2) grouping accidents and violation statistics by common driving behavior, (3) identifying subjects possibly relevant to each type of accident and violation, and (4) comparing subject scores with the number of accidents and violations of each type. (38)
4. To determine the safety effects of creating an incentive to drive safely. This can be done by informing randomly selected applicants at the outset that they will not have to take the examination when they renew their license if they drive safely during the license period. (39)
5. To determine the relationship of knowledge to safe driving. This can be accomplished by comparing examination scores to prior and/or subsequent accident histories. It can include a pass/fail comparison, but it should not be limited to whether one passes or fails. (40)
6. To determine whether providing information determined to be deficient in certain types of drivers decreases their accidents. This requires characterizing drivers and developing specialized examinations and manuals for each group. (41)



## FOOTNOTES

1. Pub. L. No. 89-54, 72 Stat. 885 (codified at 23 U.S.C. §§401 et. seq.)
2. 23 C.F.R. §1204.4 (1983).
3. Va. Code Ann. §§46.1-369 et. seq. (1980).
4. 23 C.F.R. §1204.4.
5. Va. Code Ann. §46.1-369 (1980).
6. Waller, P. F., "Do Current Licensing Procedures Screen Out Unsafe Drivers?", Traffic Safety, p. 17 (June 1976).
7. Ibid., 1977.
8. Coppin, R. S., Driver License and Driver Improvement Programs, A National Review, Australia Department of Transportation.
9. Malfetti, J. L. and Fine, J. L., "Characteristics of Safe Drivers: A Pilot Study," Traffic Safety, p. 3 (September 1962).
10. Ibid., p. 7.
11. Stoke, C.B., The Short-Term Effectiveness of Written Driver Knowledge Tests, Virginia Highway and Transportation Research Council.
12. Carpenter, D. W., 1978. The Effects of Adding Safe Driving Content to Written Law Tests Administered to Original Drivers License Applicants. California Department of Motor Vehicles and \_\_\_\_\_, 1978. The Effects of Written Licensing Tests Stressing Knowledge of Safe Driving Principles for Intermediate Record Renewal Applicants. California Department of Motor Vehicles.
13. Conley, J. A. and Huffman, W. J., 1970. "The Written Driver-Licensing Examination as a Technique in Driver Selection." Highway Research Board, Special Report No. 107, Highway Safety, p. 15.
14. Freeberg, N. E. and Creech, F. R., 1971. Development of Measures for a Driver Licensing Program in the State of North Carolina. Educational Testing Services, Princeton, N. J.
15. McKnight, A. J. and Edwards, R., "An Experimental Evaluation of Driver License Manuals and Written Tests," Accident Analysis and Prevention, Vol. 14, No. 3, p. 187. (June 1982).

16. Dreyer, D. R., 1976. An Evaluation of California's Drivers Licensing Examination. California Department of Transportation.
17. These studies are: Forbes, T. W., "The Normal Automobile Driver as a Traffic Problem." Journal of General Psychology, Vol. 20, p. 471 (1939); Burg, A., "The Stability of Driving Ability Over Time," Accident Analysis and Prevention, Vol. 2, p. 57 (1970); Soloman, D. "Highway Safety Myths." In Highway and Traffic Safety: A Problem of Definition. North Carolina Symposium on Highway Safety, University of North Carolina Highway Safety Research Center (1972); and Stewart, J. R. and Campbell, B. J., The Statistical Associations Between Past and Future Accidents and Violations, University of North Carolina Highway Safety Research Center (1972).
18. McRae, D. J., 1968. The Relation of Licensing Test Scores to Subsequent Driver Performance. University of North Carolina, Chapel Hill, N. C.
19. 402, U.S. 535 (1971).
20. Ibid., p. 539.
21. Helander, C. J., 1983. Intervention Strategies for Accident-Involved Drivers: An Experimental Evaluation of Current California Policy and Alternatives. California Department of Motor Vehicles.
22. Freeburg, N. E. and Creech, F. R., 1971. Development of Measures for a Driver Licensing Program in the State of North Carolina. Phase I: Analysis of Current Licensing Tests. Educational Testing Services, Princeton, N.J.
23. Ibid., p. 30.
24. Reel, T. O., 1970, "How Should We Be Retesting Drivers? Some Thoughts on Written Testing at Renewal," Traffic Digest and Review, Vol. 18, No. 4, p. 12.
25. Conley, J. A. and Smiley, R., 1976. Driver Licensing Tests as a Predictor of Subsequent Violations. Human Factors, Vol. 18, No. 6, p. 565.
26. McKnight, A. J. and Green, M. W., Safe Driving Knowledge Dissemination and Testing Techniques. National Public Services Research Institute (1977).
27. These examination forms are in Appendix B of each of Carpenter's studies.

28. Va. Code Ann. §46.1-369 (1980).
29. Flesch, R., How to Test Readability, Harper & Bros., N.Y.
30. Dreyer, D., 1976. An Evaluation of California's Drivers Licensing Examination, California Department of Motor Vehicles.
31. Ibid., p. 20.
32. Kelsey, J. L. and Janke, M. K., 1983. "Driver License Renewal by Mail in California", Journal of Highway Safety Research, Vol. 14, No. 2, p. 65.
33. Va. Code Ann. §46.1-380.1(e) (1980).
34. Harano, R. M. and Hubert, D. M., An Evaluation of California's "Good Driver" Incentive Program. California Department of Motor Vehicles. (1974).
35. "Why Not Use Driver License Exam to Screen for Health Problems?", Traffic Safety, Vol. 79, No. 4, p. 18.
36. See Carpenter, op. cit.
37. See Helander, op. cit. and Kelsey and Janke, op. cit.
38. See Conley and Smiley, op. cit., and McKnight and Edwards, op. cit.
39. See Harano and Hubert, op. cit., footnote 8.
40. See Freeburg and Creech, op. cit., McKnight and Edwards, op.cit., and Carpenter, op. cit.
41. See McKnight and Green, op. cit., McKnight and Edwards, op. cit., and Carpenter, op. cit.



## REFERENCES

- American Automobile Association. 1980. Sportsmanlike Driving Tests, McGraw-Hill.
- Carpenter, D. W. 1978a. The Effects of Written Licensing Tests Stressing Knowledge of Safe Driving Principles for Intermediate Record Renewal Applicants. California Department of Motor Vehicles.
- \_\_\_\_\_, 1978b. The Effects of Adding Safe Driving Content to Written Law Tests Administered to Original Drivers License Applicants. California Department of Motor Vehicles.
- Conley, J. A. and Huffman, W. J. 1970. "The Written Driver-Licensing Examination as a Technique in Driver Selection." Highway Research Board, Special Report No. 107, Highway Safety, p. 15.
- Conley, J. A. and Smiley, R., 1976. "Driver Licensing Tests as a Predictor of Subsequent Violations." Human Factors, Vol. 18, No. 6, p. 565.
- Coppin, R. S., 1977. Driver License and Driver Improvement Programs, A National Review. Australia Department of Transportation.
- Dreyer, D. R., 1976. An Evaluation of California's Drivers Licensing Examination. California Department of Transportation.
- Flesch, R., 1951. How to Test Readability, Harper & Bros., N.Y.
- Flynn, L.; Latchaw, A. J.; Ogden, A. G.; and Tannahill, W. J., 1979. Driver Licensing, Revised Ed., National Highway Traffic Safety Administration.
- Freeberg, N. E. and Creech, F. R., 1971. Development of Measures for a Driver Licensing Program in the State of North Carolina. Educational Testing Services, Princeton, N. J.
- Helander, C. J., 1983. Intervention Strategies for Accident-Involved Drivers: An Experimental Evaluation of Current California Policy and Alternatives. California Department of Motor Vehicles.
- Kelsey, J. L. and Janke, M. K., 1983. "Driver License Renewal by Mail in California," Journal of Highway Safety Research, Vol. 14, No. 2, p. 65.
- Leslie, J. D., 1971. "A Close Look at Driver Licensing," Traffic Safety, Vol. 71, No. 5, p. 16.

Malfetti, J. L. and Fine, J. L., 1962. "Characteristics of Safe Drivers: A Pilot Study," Traffic Safety, p. 3.

Marks, M. E., McNair, J. W., Jones, R. K., and Joscelyn, K. B., 1982. Identification of Motivations for Unsafe Driving Actions and Potential Countermeasures. United States Department of Transportation.

McKnight, A. J. and Edwards, R., 1982. "An Experimental Evaluation of Driver License Manuals and Written Tests," Accident Analysis and Prevention, Vol. 14, No. 3, p. 187.

McKnight, A. J. and Green, M. W., 1977. Safe Driving Knowledge Dissemination and Testing Techniques. National Public Services Research Institute.

McRae, D. J., 1968. The Relation of Licensing Test Scores to Subsequent Driver Performance. University of North Carolina, Chapel Hill.

Reel, T. O., 1970. "How Should We Be Retesting Drivers? Some Thoughts on Written Testing at Renewal," Traffic Digest and Review, Vol. 18, No. 4, p. 12.

Stoke, C. B., 1978. The Short-Term Effectiveness of Written Driver Knowledge Tests, Virginia Highway and Transportation Research Council.

Vanosall, F. E. and Rudisill, M. D., 1979. "The New Michigan Driver Performance Test," Proceedings, Twenty-third Conference of the American Association for Automotive Medicine, p. 130.

Waller, P. F., 1976. "Do Current Licensing Procedures Screen Out Unsafe Drivers?," Traffic Safety, Vol. 76, No. 6, p. 17.

"Why Not Use Driver License Exam to Screen for Health Problems?,"  
Traffic Safety, Vol. 79, No. 4, p. 18.

## APPENDIX

### STUDIES WHICH DID NOT INCLUDE THE TEST ITEMS

Carpenter, David W. The Effects of Written Licensing Tests Stressing Knowledge of Safe Driving Principles for Intermediate Record Renewal Applicants. Research and Development Section, Business and Transportation Agency, California Department of Motor Vehicles (1978).

\_\_\_\_\_. The Effects of Adding Safe Driving Content to Written Law Tests Administered to Original Drivers License Applicants. (Address same as above) (1978).

Conley, J. A. and Smiley, R. Driver Licensing Tests as a Predictor of Subsequent Violations. Department of Health and Safety Education, University of Illinois, Champaign-Urbana, Illinois (1976).

Dreyer, Dell. An Evaluation of California's Drivers Licensing Examination. Research and Statistics Section, California Department of Motor Vehicles (1976).

Freeburg, N. E. and Creech, F. R. Development of Measures for a Driver Licensing Program in the State of North Carolina, Phase I: Analysis of Current Licensing Tests. Performed by Educational Testing Services, Princeton, N. J. for the Highway Safety Research Center, University of North Carolina (1971).

McKnight, James and Green, Molly. Safe Driving Knowledge Dissemination and Testing Techniques, Volume I, General Findings. National Public Services Research Institute, 421 King St., Alexandria, Va. 22314 (1977).

