The Gerontologist
Vol. 43, Special Issue I, 29–36

Obstacles to Implementing Research Outcomes in Community Settings

Karlene Ball,[1] Virginia Wadley,[1] and Daniel Roenker[2]

Purpose: In contrast to controlled laboratory- or clinic-based research that can fail to capture the real-world behaviors of older adults, field research offers the best opportunity for ecological validity. However, the tradeoff inherent in field studies is the potential sacrifice of scientific rigor. Applied research presents a unique set of challenges that vary with context. This article discusses these challenges along with possible solutions. ***Design:*** Examples are drawn from an ongoing, longitudinal Roybal Center study of driving competence that is being conducted in Department of Motor Vehicles field sites. The challenges faced at each stage of the project are discussed. ***Method:*** Methodological issues include identifying field collaborators, approaching administrators with the research proposal, producing a battery that is manageable and acceptable while maintaining scientific merit, training indigenous personnel to administer this battery, introducing the research and consenting potential participants, and managing large data sets offsite. Additional issues include quality control, the importance of distinguishing between individuals who consent and those who decline participation, and the collection of follow-up data via telephone. ***Implications:*** The use of field research in changing public perceptions, medical practice, and public policy is discussed.

Key Words: Older drivers, Driver screening, Crash prediction

The relationships between basic research, applied research, and translational research are not always clear cut. Basic research refers to laboratory-based studies with precise control over experimental variables. Applied research can take many forms but usually involves the testing of phenomena or interventions in clinical or field site settings. Translational research involves the further step of implementing the research results in real-world settings, often under the control of an implementing agency. In these cases, the research findings may be used to make decisions about a treatment or may have some other behavioral consequence. Although basic research is traditionally thought to precede applied research, which typically precedes translational efforts, the order in which these forms of research occur is affected by multiple factors. Indeed, much basic research has arisen from the need to establish control over variables observed as a result of problem-driven or applied research.

The Roybal Centers, which are named as Centers for Research on Applied Gerontology, have focused on applied studies but have also included basic and, more recently, translational research. For example, the Roybal Center at the University of Michigan has conducted basic investigations of the relationship between age-related cognitive changes and patient behavior in medical settings. Results are being applied to the development of principles for designing medical instructions and devices. These principles will be disseminated to health care providers and medical technology developers in order to translate the findings into procedural changes. The Boston University Roybal Center has tested a fear-of-falling intervention among a patient population at risk for disablement (applied research) and has succeeded in implementing this program at other medical centers across the United States, thereby translating the research into practice.

This article reviews the history behind our own ongoing Roybal Center study of driving competence that is being conducted within the Maryland Motor Vehicle Administration, an example of translational research that stems from a decade of both basic and applied research. The article focuses on the

This research program is funded by National Institute on Aging/National Institutes of Health Grant P50-AG11684 (Edward R. Roybal Center for Research on Applied Gerontology).

Address correspondence to Karlene Ball, PhD, Department of Psychology, Edward R. Roybal Center for Research on Applied Gerontology, University of Alabama at Birmingham, HM 100, 1530 3rd Avenue South, Birmingham, AL 35294-2100. E-mail: kball@uab.edu
[1]Department of Psychology and Edward R. Roybal Center for Research on Applied Gerontology, University of Alabama at Birmingham, Birmingham, AL.
[2]Department of Psychology, Western Kentucky University, Bowling Green, KY.

challenges faced in moving this research from the laboratory to the real world.

Background

One of the guiding research questions for the University of Alabama at Birmingham (UAB) Roybal Center has been "Why are older adults, as a group, overly represented in motor vehicle collisions and particularly injurious crashes?" Although this question has been investigated for some time in the literature, our interest in this area grew from our discussions with older adults who participated in our laboratory-based studies on vision and aging. In 1979, as the result of some research results in this basic area, we began to work on developing new methods for evaluating the everyday vision and cognition of older adults. For many of these studies, the questions revolved around identifying the underlying mechanisms for everyday visual problems reported to us by our older participants, as well as determining whether interventions for declines in these mechanisms were feasible. One application that developed from this basic research was a new measurement tool dubbed "the Useful Field of View" (Ball, Beard, Roenker, Miller, & Griggs, 1988). Then, during the 1980s and 1990s we conducted concomitant basic research aimed at understanding why older adults had increased difficulty with this vision-based test of processing speed and attention skills, as well as understanding what accounted for improvement in performance on the test with training. These efforts are described below.

Asking Applied Research Questions

As a result of some questioning from National Institute on Aging (NIA) program staff regarding whether the Useful Field of View test might be related to driving competence, a series of studies was launched investigating this question. In one of the earlier studies, 306 older adults were recruited for participation in a case-control study involving a 5-hr assessment battery, which was carried out in the Eye Foundation Hospital at UAB. This very comprehensive battery evaluated eye health, visual sensory function, visual function, cognitive function, physical function, medications, medical diagnoses, and driving habits, among other things (Ball, Owsley, Sloane, Roenker, & Bruni, 1993).

Identifying Early Field Collaborators

Importantly, the above study enlisted the cooperation of the Alabama Department of Public Safety (DPS) as one of our earliest field collaborators. Through an initial contact of Dr. Cynthia Owsley with the Alabama DPS Medical Advisory Board, further meetings were held to discuss the feasibility of research incorporating crash records from the state of Alabama as an outcome measure for our research. This early collaborative link enabled us to shift toward an applied research focus; specifically, the collaboration allowed us to establish relationships between our research battery and driving competence, verified through crash records, and led to several publications in this area (e.g., Ball et al., 1993; Goode et al., 1998; Owsley et al., 1998; Sims, McGwin, Allman, Ball, & Owsley, 2000).

Expanding to Field Research

The theme of our Roybal Center, established in 1993, is "Enhancing Mobility in the Elderly." Initial research studies of our Center for Research on Applied Gerontology focused on the evaluation of both a visual intervention (cataract surgery) and a cognitive intervention (speed of processing training) across a variety of mobility outcome measures, including driving competence. The center provided resources that were instrumental in expanding our list of collaborators and in moving the research endeavors from the laboratory to a variety of field sites. For example, as new measurement tools were developed, we were able to make them available to other investigators involved in similar research areas. The center was able to provide Useful Field of View technology; a variety of validated questionnaires on driving habits, falls, and other mobility measures; a Timed Instrumental Activities of Daily Living kit; and ultimately a driving simulation facility for a number of collaborative studies on campus. This ability to share our resources resulted in collaborative agreements with the Salisbury Eye Study at Johns Hopkins University, a study on Alzheimer's disease and driving at Washington University, a study on Heads Up Technology with Scientex, Inc., a study on driving competence at the University of Iowa, and a population-based study within the Maryland Motor Vehicle Administration (MVA).

The importance of giving presentations and placing peer-reviewed publications in diverse outlets should also be mentioned as a key ingredient in helping to form collaborations and contacts with outside businesses and agencies involved in driving research. Presentations to mixed groups of researchers and service personnel, such as the audience in attendance at meetings of the Transportation Research Board and its Committee on Safety and Mobility of Older Drivers, were instrumental in establishing contacts that resulted in collaborations on the Maryland project. Publications from our Roybal Center, and the resulting national media coverage, also led to numerous contacts with individuals in organizations such as AARP, law enforcement, other state departments of motor vehicles, the American Automobile Association, hospitals, clinics, driving rehabilitation services, state highway administrators, senior centers, insurance companies, automobile manufacturers, and other companies interested for a variety of reasons in the assessment of older persons. These contacts led to invitations to make presentations at meetings that would not typically be attended by members of the scientific community, but which provide an opportunity to establish further contacts for applied research venues.

As an outgrowth of contacts such as these, field collaborators with common interests in the safety and mobility of older drivers were identified. Specifically, Dr. Robert Raleigh of the Maryland MVA and Dr. Loren Staplin of Transanalytics, Incorporated (funded through a contract from the National Highway Traffic Safety Administration, or NHTSA), joined forces with investigators from UAB and Western Kentucky University to conduct a large-scale study of a brief screening battery designed to predict crash and other mobility outcomes among older drivers in the state of Maryland. Thus, this translational field study was jointly funded by NIA (through the UAB Roybal Center), NHTSA, and the Maryland MVA. The study has screened over 5,000 participants, enrolled 2,207, conducted first annual follow-up interviews with a subset of the sample, and is in the process of conducting second annual follow-up interviews. Preliminary study results with respect to crash risk and other mobility outcomes have been reported at national scientific and other professional meetings, and a paper reporting these outcomes will soon be submitted for scientific review (Ball, Roenker, et al., 2002).

Developing the Study Battery

Any test battery to be implemented in a field setting must balance the concern for scientific merit with practical considerations such as whether the protocol will be acceptable to the target population of older adults and whether it will be feasible to administer outside the laboratory and, as is often the case, by nonscientist test administrators. In the study being conducted in the Maryland MVA, for example, our goals were to compile a screening battery that would (a) effectively identify older drivers at risk for crash involvement, (b) be palatable—and nonthreatening—to the target population of older drivers, (c) be acceptable to MVA-based test administrators, and (d) be manageable and portable for use in multiple settings and for possible future use by health care professionals. In addition to these goals, we wished to maintain scientific integrity in both the selection and the implementation of the battery.

The first step for the study investigators was to become acquainted with existing science concerning performance-based predictors of crash risk among older adults. Our own research over the previous decade informed the battery selection process, along with a thorough search of pertinent scientific and medical literature. In essence, the team of study collaborators compiled and discussed potential battery components, weighing each with respect to its reliability and validity, crash prediction potential, availability of age-appropriate normative data, ease of administration, and, perhaps above all, the cost and benefit yielded for each minute of test time. Indeed, the need to keep the battery brief was the overriding principle dictating which tests were ultimately included in the battery. In some cases, brief subtests of more extensive instruments were selected for inclusion in the final screening battery.

We relied on existing measures with known psychometric properties in the Maryland study because this strategy allowed us to compare results obtained in this setting with results obtained under controlled testing conditions. The process of developing, piloting, and validating new or modified instruments was beyond the scope and resources of the study.

General considerations in battery selection or development include the incorporation of multimethod, multimeasure procedures to the extent possible within a given field setting, in order to avoid test-specific effects that are difficult to interpret. An additional consideration in developing the protocol is its vulnerability to order effects such as practice and fatigue that can affect performance on individual components of the battery and obscure results with respect to the battery as a whole.

In the Maryland study, the final screening battery that met the criteria discussed above was a 15-min assessment composed of performance-based and self-report measures of medical conditions, physical function, visual function, and cognitive function. Specifically, the battery was composed of the Useful Field of View (UFOV) Subtest 2 (Ball & Owsley, 1993); the Gross Impairments Screening Battery (Staplin, Lococo, Gish, & Decina, 2002), incorporating existing measures of cognitive function and physical function; and a questionnaire assessing health, medication usage, driving habits, and general mobility.

Training Field Personnel

Numerous problems must be solved before a measurement tool is moved to a real-world setting. It is one thing to develop a measure that will be administered by research-savvy individuals in a research setting to relatively test-wise participants and quite another to administer that same test using novice test administrators testing the wary public in a noisy and uncontrolled public environment. Much thought and preparation must take place before a test can be moved from one environment to the other. First, the essential parameters of the test and testing technique must be identified and the nonessential elements discarded. For example, with the UFOV test, viewing distance from the monitor is controlled when the test is administered in the laboratory setting, but control over the individual's head position is not likely to take place in a field setting. Second, the essential elements of the test need to be simplified and standardized so as to make administration of the test as free of the role of the administrator as possible. Again, in the case of the UFOV measure, years of experience taught us how to simplify the test and the test instructions and how to present them on the computer monitor so that only the critically necessary information was presented. When the UFOV test was first developed, it required a highly trained individual to administer the test, and the test protocol lasted nearly 45 min. The version of the test used in this project can be self-administered, at least with some segments of the population, and requires only 3–5 min to complete.

Our experience with automated testing led to a broader discussion of the influence of test administrators on successful data collection efforts. Must a protocol be completely computerized, or can tests be used that require some degree of judgment of the administrator? In general, the more automated that test instructions, stimulus materials, test protocol, test scoring, and data storage can become, the better. Nevertheless, the conclusion reached by our research group was that nonautomated tests requiring facilitation and explanation by the administrator can be feasible in an applied setting given careful test selection and tester training. In summary, substantial time and energy were devoted to the choice and refinement of each measure before any attempt was made to move into the field setting and train the field administrators.

Once testing technique has been refined, there are a number of obstacles to moving the polished test from the laboratory to the field. Probably the single largest problem is achieving consistency of test administration. During earlier stages of test development, we gave little thought to the participant–test administrator interaction because our test administrators were either ourselves or highly trained students and staff. The rigor of the scientific method and the need for consistency in the assessment process are part of our daily experience. These are skills that lie at the core of a research staff. It is unlikely that novice administrators identified in a field setting will have the same respect for scientific rigor that we take for granted in laboratory staff. Although potential test administrators can be informed of the need for such rigor, it is unlikely that such information will become an ingrained part of the test administration. It has been our experience that novice testers will modify the testing procedure over time, usually in ways that they believe make test administration easier for them. Changes in protocol most likely will be subtle and cumulative, thereby representing a slow evolution of the test procedure over time. Such changes may be important, such as changes in the content of instructions, or trivial. In either case, the testing procedures must be adapted in advance with the goal of minimizing the impact of the test administrators' potential to drift from protocol. This can be achieved by giving some thought to how staff will be trained. For example, we found videotape to be a useful tool for modeling proper test administration techniques. This tool shows correct performance of the test and can also serve as a reference source against which to point out procedural "evolutions." Second, we used unannounced quality control visits as an effective method for discovering changes in test administration. Not only were we able to use these visits to restore proper administration practices, but they also served to reemphasize the importance of the data being collected. Third, we used computer automation to minimize the role of the test administrator to the extent possible.

Finally, some thought was given to the personal characteristics of test administrators. Because access to personnel in a field setting is generally limited, it is likely that these same individuals will serve in a recruitment role in addition to a test administrator role. Therefore, their style of interaction with the general public is an important consideration.

Approaching and Consenting Participants in the Field

Approaching potential study participants in a given field setting calls for creative strategies that minimize possible defensiveness or even alarm. The recruitment approach should minimize any obstacles to participating, maximize the benefits of participating, and capitalize on the existing skills of field staff. In large part, the approach will be dictated by the particular setting. In the Maryland study, we sought to recruit older drivers in a MVA setting. The advantage of this setting was that it was the ideal place to capture all older drivers presenting for license renewal. Thus, it was both a sensible and feasible setting for the research. The disadvantage, however, was that it was an unexpected and less credible institution within which to conduct research than a university setting or a physician's office. As such, the setting itself presented a potential threat to acceptance of the research by the study population. Furthermore, the MVA setting is vested with the authority to grant or suspend driving privileges, and this association in the mind of the older driver might have presented a powerful obstacle to successful recruitment without attention to this possibility at the outset. As part of the study design, therefore, individuals were approached for study participation only *after* renewing their drivers' licenses. In addition, potential participants were explicitly assured that their performance on the screening battery under evaluation would *not* in any way affect their driving privileges. This strategy did not eliminate less proficient drivers but did limit the sample to those with adequate vision, as the only criterion for license renewal was a threshold level of visual acuity (20/40).

Our attempts to minimize other obstacles to participation in the Maryland study included keeping the testing session as brief as possible and considering—and then trying to counter—other costs and factors that might prevent an individual from participating (e.g., perception of the study as a hassle). At the same time, we attempted to maximize the benefits of participating in the study. Individuals who agreed to participate were allowed to move to the front of the waiting line for procurement of their license, for example, as a personally relevant incentive. They were also presented with a rationale for the study, highlighting its importance from a perspective of public safety and its potential to enhance safe mobility if successfully implemented on a larger scale in the future. In this way, an appeal to altruistic motives was made.

We also attempted to capitalize on staff strengths in the area of recruitment by selecting, to the extent possible, individuals who could rapidly establish rapport with older adults and competently guide them through the informed consent process. A combination of strategies therefore provided the framework for introducing the field study and carrying it out in a unique setting.

Characterizing Consenters Versus Decliners

As discussed, the recruitment of participants in a field setting poses somewhat different problems than traditional laboratory recruitment issues. One of the most important considerations is that the enrolled sample must mirror the general population. Although there are numerous texts on the statistical and experimental design issues of participant selection, the issue at hand is the practical problems associated with field recruitment of participants. Clearly, the field environment presents the opportunity for sampling biases to operate. In the case of the Maryland project, MVA staff recruited participants from license renewal applicants, and the situation did not permit an orderly approach to selection of participants from the stream of licensees. Recruiters selected individuals as space and the normal work schedule allowed. Biases introduced by appearance, gender, or other characteristics may have operated and gone undetected. The only way to evaluate the extent to which such biases may have operated was to have a method for checking the characteristics of the participants against those who declined participation. Such information is available only if there are public records available (e.g., license renewal board) or if the participants are recruited from a known database. In either case, the ability to demonstrate the similarity of those who participate to those who do not is critical to the validity of the research. Fortunately, we were able to successfully demonstrate in the Maryland study that participants did not differ from decliners on any demographic variables or in crash history. In the event that such databases are not available to researchers in other domains, an exit interview, albeit brief, may generate the relevant demographic information on decliners.

Collection of Follow-Up Telephone Data

As part of the Maryland project, we also contacted a subsample of the participants by telephone and had them complete an interview about their driving skills, general mobility, health, medications, and related information. Telephone interviewing presents a unique set of problems, particularly if the sample consists of older adults. Because the interview process is somewhat lengthy, we developed techniques to limit attrition. In particular, we encouraged our interviewers to attend to and record small bits of personal information offered by participants (e.g., recent births of grandchildren, sicknesses, travel plans, etc.). In the process of doing so, the interviewers are able to establish rapport with the individual, which facilitates the interview process. At the same time, we cautioned our interviewers against letting this information alter the manner in which the interview data were collected, so that the scientific integrity of the project is maintained. Nevertheless, these additional bits of personal information have allowed the interviewer to return to the participant at a 1-year interval and reestablish the rapport developed during the initial interview. This method has helped us to overcome the reticence that many older adults feel about releasing information to unknown parties. In fact, the technique has been so successful that in some cases the 30-min interview has required an hour or more to complete. The effectiveness of this simple technique is reflected in the fact that the 1-year attrition rate for the project was close to 5%, a remarkable figure given the initial mean age of the sample (70 years).

Managing the Data

With data collection procedures in place, the problem of managing the field data has presented another challenge to the Maryland project investigators. First, although the investigators agreed that the data should be stored on site in Maryland, each site wanted to maintain a copy of the database and to perform independent analyses. Because data checking, corrections, modifications, and coding for analyses have occurred regularly across sites, any change that occurs at one site has the potential to result in loss of a common database unless strict procedures are observed for multisite notification of data "cleaning" and coding processes and outcomes. It has been, therefore, no small challenge to maintain a truly common database.

Furthermore, the collaborators on this project represent a variety of scientific disciplines as well as the fields of medicine, industry, and government. Consequently, the amount and nature of prior training in the areas of data management and statistics among this group is diverse. More to the point, each stakeholder in the project has had a different investment in the study outcomes. Not surprisingly, then, a great deal of discussion has ensued as to the appropriate site for maintenance of the data, the appropriate analytic approaches to the data, the appropriate cut points to use for each test in the screening battery in order to maximize sensitivity and specificity, and the appropriate interpretation of the study results. Real-world challenges have included lively discussions relative to sharing data, questioning of biases associated with each stakeholder's perspective, and questioning of statistical results and conclusions based on them. Nevertheless, these differences and challenges have made for a fruitful project in which no one investigator is able to dictate a "right" way to handle the data. In some respects, the compromises reached resemble the peer review process that can only serve to enhance the enterprise of research.

Changing Medical Practice, Public Perception, and Public Policy

The history of research assessing the UFOV and its relationship to driving competence provides a good illustration that research can accomplish both theoretical and applied goals. On the basis of more than 50 published studies, as well as preliminary data from the large-scale Maryland study, we have converging evidence that UFOV is highly predictive of crash involvement in older drivers. Indeed, recent meta-analytic results indicate that the impact of UFOV impairment on crash risk is very high across studies (mean effect size $D = 1.195$), the impact of visual acuity is very low ($D = 0.079$), and the impact of mental status is moderate (D ranging from 0.539 to 0.632; Ball, Wadley, et al., 2002). Thus, UFOV performance is a much stronger predictor of crash risk than many factors that have been more traditionally examined.

There is also growing evidence that speed of processing training can result in both improved UFOV and improved driving skills. Ongoing studies are examining the effects of speed of processing training, relative to various types of sham training, on driving simulator performance, driving habits, on-road performance in an instrumented car, and other mobility outcomes. Preliminary findings suggest that speed of processing training not only improves UFOV performance but may transfer to other cognitive functions as well as to certain aspects of driving (Edwards et al., 2002; Roenker, Cissell, Ball, Wadley, & Edwards, in press; Rumble, Edwards, Clay, Woodley, & Ball, 2001).

There are, however, numerous obstacles encountered in attempting to take research results such as these and infuse them into public policy or medical practice. In the case of implementing research findings at the level of medical practice, it can be difficult to quantify the associated costs and benefits. For example, what are the health care costs associated with mobility loss and to whom do those costs accrue? How much money does continued mobility save the health care industry, and what cost is incurred for training programs that result in this savings benefit?

Translating research outcomes into changes at the level of legislation is a slow process. Once research outcomes indicate that a change in policy is needed, an institution such as the MVA may have procedures in place to prepare legislative proposals for consideration by the state assembly. First, the nature of these proposals must be determined. For example, the MVA might consider the benefits of a mandatory testing program for older drivers versus a recommendation for further use of voluntary testing programs. Once the content is decided and the legislative proposal is drafted, adoption of the legislation tends to occur only after years of well-publicized research findings and promotion of the legislation by interested parties—parties with funds. Thus, the process of instituting various licensing procedures, including adoption of more sensitive screening tests, is difficult and time consuming and is made even more difficult due to the fact that in many cases changes must occur state by state rather than at a national level. Each state prefers to do its own driving evaluation, and the most popular way of conducting that evaluation is of course "the way it's always been done here." Most states resist being the first to adopt something new. Thus, there appears to be a natural resistance to change that works against implementation of new programs, despite their scientific basis.

Another obstacle to implementation of study findings is that the implementing agency may be afraid of the public's perception of the change. Thus, it may be necessary to work on changing public perceptions before instituting policy change. In some cases, collaborations with various media outlets to establish a public information campaign may be helpful.

In the special case of driver evaluation, acceptability of driver testing for the driver him- or herself is a related issue. Again, the acceptability of an evaluation will be contingent on what is at stake given the outcome of that evaluation. There are a number of potential strategies available to minimize the impact of visual, cognitive, and physical impairments on driver safety among older adults (e.g., treatment of eye disease, cognitive training programs, physical therapy). If the outcome of evaluation and remediation is that the length of time that an older driver may continue to drive safely is maximized, then the initial evaluation may be viewed in a more positive light.

To address this very complex public health issue, a comprehensive approach is needed. Currently, mandatory driver retesting is conducted in many states for those drivers who are referred to a medical advisory board due to concern on the part of family and friends, poor driving records, or a particular medical diagnosis. This approach should not be abandoned, but it could be improved. One concern with respect to any mandatory retesting program is that it may lead some potentially fit older drivers to stop driving rather than submit to retesting. Therefore, as discussed above, any such program needs to be accompanied by a strategic plan for informing the public of the testing procedures to be used and of the remediation options that will be made available to those with impairments that are amenable to rehabilitation.

A controversial question that has been widely debated is whether age alone should trigger a mandatory driving evaluation. Mandatory retesting for functional impairment at licensing agencies would obviously be quite costly if all licensed drivers were retested. Thus, discussions have turned to whether retesting could or should be limited to only those drivers over a certain age. For example, legislation was introduced in June 1999 in the California State Assembly that would have required all drivers aged 75 and above to demonstrate driving proficiency through both written and road tests in order to renew their driver's licenses. This proposed legislation did not pass, and the proposed trigger for reexamination was changed from an age criterion to a criterion based on prior crashes, violations, or referral for evaluation.

One perspective on this question is that the reexamination of functional abilities needed to drive safely should be viewed as a form of preventive medicine. According to this view, the occurrence of functional decline occurs with a much higher prevalence in older age groups, so screening for such decline should not begin until older age, just as screening for certain medical conditions in the absence of any symptoms is routinely based on attaining the age at which risk increases. This approach would allow for baseline measurement of individual performance against which any indication of decline could be noted at a future date.

A specific example of an age-related functional decline that is predictive of driving outcomes is decline in the useful field of view. Both UFOV reduction and crashes are more prevalent with increasing age (Ball, Roenker, & Bruni, 1990; Transportation Research Board, 1988); however, UFOV reduction is substantially better than chronological age at distinguishing between drivers who are at risk for crash and those who are not (Ball et al., 1993). The probability of a hit (accurate prediction of crash) versus a false alarm (inaccurate prediction of crash) for independent predictor variables was calculated by Ball and colleagues (1993). Results demonstrated that the UFOV test ($d' = 2.27$) was superior not only to age ($d' = 0.58$) but to vision parameters (d's ranging from 0.24 for acuity to 0.67 for contrast sensitivity) and mental status tests ($d' = 0.59$) in accurate prediction of crash, with sensitivity of 89% and specificity of 81%. Thus, if an individual was crash-involved, the likelihood of his or her having a UFOV reduction greater than 40% was 0.89, whereas the likelihood of an individual who was not crash-involved having intact UFOV ($\leq$ 40% reduction) was 0.81.

This illustration of accounting for age-related functional declines versus age alone takes on additional importance when one considers the implications of misclassifying drivers. What happens to drivers wrongly classified as impaired or high risk for crash when they are not? Individuals' autonomy and quality of life could be jeopardized by such misclassifications if they led to automatic revocation of the driver's license. Therefore, a classification of a driver as high risk should lead to further assessment, with driving outcomes contingent on the results of multiple performance indices. Conversely, what are the ramifications of wrongly classifying drivers as safe when they are impaired? In these cases, the individuals' safety, as well as public safety, are at stake. Thus, errors made in either direction could prove costly both for the individual and for society.

The challenges of implementing research outcomes such as those described here into community practice, state law, or public policy are formidable. Much groundwork has been laid—a process that has taken time, effort, and considerable resources. In our quest to disseminate research on older drivers and translate the results into policy-level changes with the potential to improve quality of life through sustained mobility, we have tried to establish as many collaborations as possible with industry, government, and older adult advocacy groups. We have recently sought funding in partnership with small business to develop educational videotapes for physicians and other health care providers in order to more directly inform this group of the older driver research and engage them in the practice of evaluating and intervening on behalf of high-risk patients. The Roybal Center program provided the impetus and the dollars for our evolution from basic to applied and translational research. It is our hope that this investment will ultimately reap benefits for older adults by enhancing their safe mobility and extending their independence for as long as possible.

References

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. *Journal of the Optical Society of America, 5*, 2210–2219.

Ball, K., & Owsley, C. (1993). The useful field of view test: A new technique for evaluating age-related declines in visual function. *Journal of the American Optometric Association, 64*(1), 71–79.

Ball, K., Owsley, C., Sloane, M., Roenker, D., & Bruni, J. (1993). Visual attention problems as a predictor of vehicle crashes in older drivers. *Investigative Ophthalmology & Visual Science, 34*, 3110–3123.

Ball, K., Roenker, D. L., & Bruni, J. R. (1990). Developmental changes in attention and visual search throughout adulthood. In J. Enns, (Ed.), *Advances in psychology* (Vol. 69, pp. 489–508). North-Holland: Elsevier Science.

Ball, K., Roenker, D., McGwin, G., Wadley, V., Edwards, J., Raleigh, R., et al. (2002). Can high-risk older drivers be identified through performance-based measures in a Department of Motor Vehicles setting? Unpublished manuscript.

Ball, K. K., Wadley, V. G., Edwards, J. D., Roenker, D. L., Frankel, M. S., Myers, R., et al. (2002). UFOV as a predictor of everyday mobility: Review and meta analysis. Unpublished manuscript.

Edwards, J. D., Wadley, V. G., Myers, R. S., Roenker, D. L., Cissell, G. M., & Ball, K. K. (2002). The transfer of a speed of processing intervention to near and far cognitive functions. *Gerontology, 48*, 329–340.

Goode, K., & Ball, K. (1998). Useful field of view and other neurocognitive indicators of crash risk in older adults. *Journal of Clinical Psychology in Medical Settings, 5*, 425–440.

Owsley, C., Ball, K., McGwin, G. Jr., Sloane, M. E., Roenker, D. L., White, M. F., et al. (1998). Visual processing impairment and risk of motor vehicle crash among older adults. *Journal of the American Medical Association, 279*, 1083–1088.

Roenker, D. L., Cissell, G. M., Ball, K., Wadley, V., & Edwards, J. (in press). The effects of speed of processing and driving simulator training on driving performance. *Human Factors.*

Rumble, M., Edwards, D., Clay, O., Woodley, R., & Ball, K. (2001, August). *ACCELERATE: A cognitive intervention to improve everyday mobility.* Paper presented at the Annual Meeting of the American Psychological Association, San Francisco, California.

Sims, R. V., McGwin, G., Allman, R. M., Ball, K., & Owsley, C. (2000). Exploratory study of incident vehicle crashes among older drivers. *Journal of Gerontology: Medical Sciences, 55A*, M22–M27.

Staplin, L., Lococo, K. H., Gish, K. W., & Decina, L. E. (2002). *Model driver screening and evaluation program final technical report* (DOT-HS-XXX-XXX). Washington, DC: National Highway Traffic Safety Administration.

Transportation Research Board. (1988). *Transportation in an Aging Society* (Vol. 1). Washington, DC: National Research Council.

Received July 15, 2002
Accepted September 16, 2002
Decision Editor: Richard Schulz, PhD