IDEA

Innovations Deserving
Exploratory Analysis Programs

***Safety IDEA Program***

# Assessment of driver safety in commercial motor vehicles

Final Report for
Safety IDEA Project 05

Prepared by:
Michael B. Cantor, Ph. D.
WayPoint Research, Inc.
Atlanta, GA

***September 2006***

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

**INNOVATIONS DESERVING EXPLORATORY ANALYSIS (IDEA) PROGRAMS MANAGED BY THE TRANSPORTATION RESEARCH BOARD**

This Safety IDEA project was funded by the Safety IDEA Program, which focuses on innovative approaches for improving railroad safety and intercity bus and truck safety. The Safety IDEA Program is funded by the Federal Motor Carrier Safety Administration (FMCSA) and the Federal Railroad Administration (FRA).

The Safety IDEA Program is one of four IDEA programs managed by TRB. The other IDEA programs are listed below.

The Transit IDEA Program, which supports development and testing of innovative concepts and methods for advancing transit practice, is funded by the Federal Transit Administration (FTA) as part of the Transit Cooperative Research Program (TCRP).
The NCHRP Highway IDEA Program, which focuses on advances in the design, construction, and maintenance of highway systems, is funded as part of the National Cooperative Highway Research Program (NCHRP).
The High-Speed Rail IDEA Program, which focuses on innovative methods and technology in support of the next-generation high-speed rail technology development, is funded by the FRA.

Management of the four IDEA programs is coordinated to promote the development and testing of innovative concepts, methods, and technologies for these areas of surface transportation.

For information on the IDEA programs, look on the Internet at www.trb.org/idea, or contact the IDEA programs office by telephone at (202) 334-3310 or by fax at (202) 334-3471.

IDEA Programs
Transportation Research Board
500 Fifth Street, NW
Washington, DC 20001

# *ASSESSMENT OF DRIVER SAFETY IN COMMERCIAL MOTOR VEHICLES*

**Final Report**

Safety IDEA Project 05

Prepared for

Safety IDEA Program

Transportation Research Board

National Research Council

Prepared by

Michael B. Cantor, Ph. D.

WayPoint Research, Inc.

Atlanta, GA

September 2006

ACKNOWLEDGEMENTS

I wish to express my thanks to colleagues who were especially helpful in contributing to this Safety IDEA project.

Richard Anderson, Director of Driver Training at Baltimore Gas and Electric, made many insightful suggestions on how WayPoint measures can be translated into practical driver training strategies.

Donald Lane, an owner-operator OTR driver and master trainer, has been a mainstay in linking WayPoint concepts to what long-haul drivers deal with every day. He has also been a valuable colleague, lending his wide-ranging technical expertise to our understanding of the data.

Kurt Wiesenfeld, Ph. D., Professor of Physics at Georgia Tech, joined in many discussions of how WayPoint data may or may not be interpretable as a non-linear dynamical system.

Marc Stieglitz, Ph. D., Associate Professor of Civil Engineering at Georgia Tech, also addressed the interpretation of the data in terms of non-linear dynamics in many helpful and delightful discussions.

Finally, I would like to express special appreciation to Harvey Berlin, Senior Program Officer, Transportation Research Board, for his thoughtful guidance throughout this Safety IDEA project.

TABLE OF CONTENTS

LIST OF FIGURES

## Executive Summary

This report describes the results and conclusions of *Assessment of Driver Safety in Commercial Motor Vehicles*, a Safety IDEA project aimed at providing further validation of WayPoint (WP), a 4-minute, non-verbal, web-based assessment of a truck driver's propensity for having preventable collisions. Key applications of the WP assessment are: a) to select safer driver/applicants; b) to determine who among existing drivers would benefit most from training; c) to diagnose which issues should be emphasized in training.

A total of 1,218 truck drivers from seven different fleet operators completed WP as part of this project. Criterion measures were collected from each of the companies and included, at a minimum tenure with the company and number of preventable ("at fault") and non-preventable collisions. The primary criterion measure was based on a driver's preventable crash frequency adjusted for exposure and expressed as an *odds ratio*, i.e. a driver's *crash frequency/average crash frequency in the population of all drivers in the data base. An odds ratio of 1.0, for example, indicates that a driver's crash rate is average among all drivers; 3.0, three times the average, and so on. Four levels of a* Poisson-distributed odds ratio were discernable: a) 0.5, half of all drivers, b) 1.2, 35% of all drivers, c) 3.4, 11% of the drivers, and d) 6.2, 4% of the drivers. In 79 cases, we were provided with a record of complaints from the driving public and the cost of each preventable crash.

WP identified 81% of drivers whose crash risk was six times average and 69% of drivers whose risk was three times average. At the same time, the false positive rate, categorizing a driver as "high risk" (three or six times average risk) when, *in fact*, he or she was "low risk" (less than three times average risk) was only 8%. This degree of sensitivity is unusually high by testing standards and applies to the full sample. Development of the scoring algorithm continues.

The WayPoint scoring algorithm is based on a new approach to behavioral analysis. Using information theory, four derived performance measures, i.e. sources of variance, were combined to predict crash proneness, including: a) *Flow Uncertainty*, the degree to which driving speed, as predicted by the WP "channel capacity" measure, deviates from the population average and b) *Driver's Uncertainty*, the degree to which the individual's moment-to-moment speed deviates from his baseline. The latter is based on the three WP "situational awareness" measures: impulsivity, sustained attention, and distractibility.

Among the key findings under this project was the discovery that a particular WP profile includes only 24% of all OTR drivers yet accounts for a conservative 62% of drivers with the highest crash risk.

The IDEA program was critical to the development of the WP product. Because the algorithm had to be accurate enough to identify individual crash-prone drivers rather than providing an average risk based on an average profile, we tried many different quantitative strategies. The behavioral science turned out to be completely new. This IDEA project gave us the resources and time to find a workable solution to a very complex problem. As a small business, we could not have succeeded without it.

Deployment of WP to the trucking industry was begun during the course of this project. An Eastern gas/electric utility has been using the WP assessment to diagnose the training needs of their drivers and has had it evaluated by a university-based transportation institute; the user has indicated that the evaluation found WP to be unusually sensitive.

Given the chronic shortage of truck drivers, especially long-haul, truck-load drivers, carriers are wary of anything that might reduce their pool of drivers. On the other hand, given the cost of collisions, the expense of unnecessary training, and the need to pinpoint training opportunities, we believe that carriers will be open to four-minute WP, especially after a blind trial that correctly identifies their drivers with a crash history while not false-alarming on their safe drivers.

## IDEA PRODUCT

WayPoint (WP) answers a need for a quick, highly accurate, low-cost assessment of a driver's aptitude to operate a commercial motor vehicle. The WP assessment is administered on the internet, takes four (4) minutes from start to finish, and requires no reading beyond a few preliminary instructional sentences.

After a driver takes WP, his or her management is provided with an odds ratio that compares the driver's risk of a preventable collision relative to that of other drivers on the road. Both driver and management receive a feedback report that features a) the driver's typical speed relative to other drivers, b) his or her "situational awareness" regarding key aspects of the visual field and, most importantly, c) how the two variables interact, thereby determining the chances of a preventable collision. As applicable, the report makes specific suggestions on how the driver, along with driver trainers, can reduce the chance of a preventable crash. WP helps commercial fleet operators to:

a) **Select safer driver/applicants while expanding the pool**. Currently, commercial drivers must be at least 21, though the FMCSA has considered (and rejected) a minimum age of 18 to address the driver shortage. WP identifies collision-prone drivers without reference to age, potentially expanding the pool of safe drivers. Virtually all companies interview driver candidates despite unevenness in interview quality. Some companies attempt to overcome the subjective interview with a personality survey, though such tests are time-consuming and notoriously low in validity. At best, they identify the few crash-prone applicants at the extremes of personality.[i] Most companies order an MVR, though three years is usually too short a time to assess *an individual's* tendency to crash, given the low base rate of crashes overall. An MVR, of course, is not applicable for inexperienced driver candidates.

b) **Identify existing drivers who would benefit most from training.** Only a fraction of drivers benefit from training; the great majority know very well how to avoid collisions. WP can identify those who would benefit most from training, allowing companies to either save money over the "train everybody" approach or put the extra dollars into *better* training, e.g. driving simulators, for the most needy.

c) **Diagnose a driver's specific training needs.** Training focused on an individual's weaknesses has a much better chance of reducing collisions than generic training. According to our data, for example, not all drivers should be told to *expand* their field of view ("big picture"). Indeed, some should *narrow* their visual field to better match their typical driving speed.

"You can't improve what you can't measure." This truism is particularly true in the area of safety and human performance. Through better measurement, WP improves fleet safety by identifying crash-prone drivers *before* they are hired and by helping companies focus better training on existing drivers who need it most.

## WAYPOINT CONCEPT

WP consists of four screens, each randomly scattered with fifteen letters and numbers. The user's task is to quickly move the cursor from number to letter in sequence: 1 A 2 B 3 C, etc. On the fourth screen, distracting icons are interspersed among the letters and numbers. Two WP measures interact to predict a driver's likelihood to have preventable collisions, speed and variation in speed.

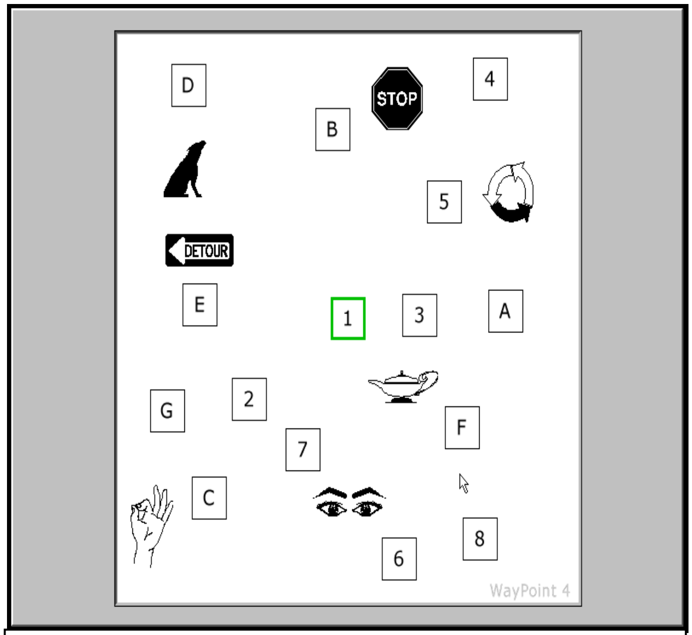


**Figure 1** Fourth screen of WayPoint assessment showing distracting icons among the numbers and letters

SPEED

The raw data consist of four speeds, i.e. letter-number items *correctly* touched per second.

Previous research found that speed on WP was significantly correlated with driving speed, crash frequency, and frequency of moving violations (Figure 2 a, b). Speed on WP also correlates with reading speed, speed of performing arithmetic calculations, and even the speed at which pharmacists fill prescriptions. [ii]

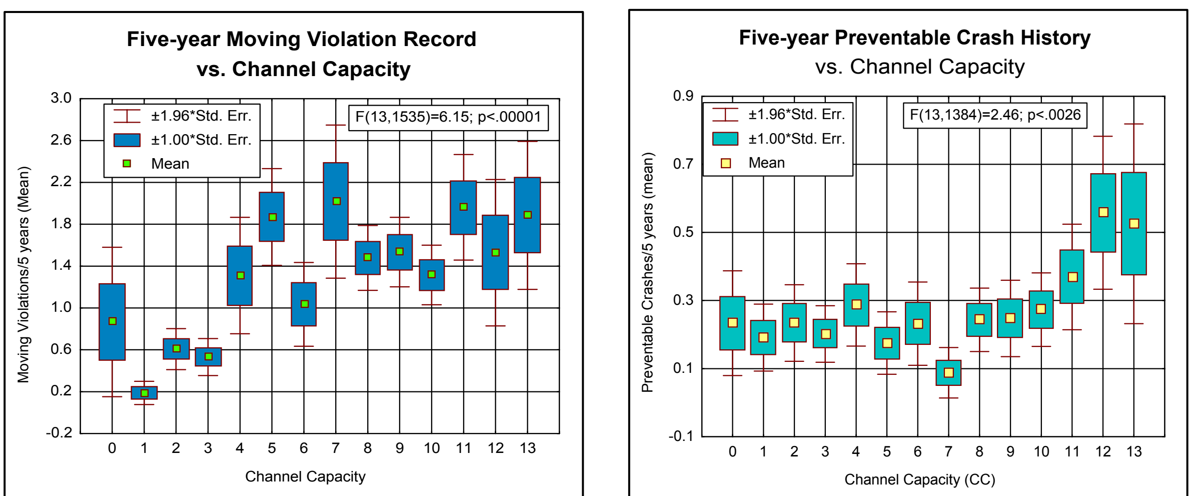


**Figure 2 a, b** As CC increases in steps of approximately one "just noticeable difference", mean 5-year moving violation frequency increases linearly overall in zigzag (see below) fashion. Crash frequency is also positively correlated with CC. The greater frequency of violations vs. crashes would account for the lack of a clear zigzag pattern of crashes.

Given the correlations between speed on WP and many different behaviors, speed can be seen as a sensitive measure of a person's information throughput---his bandwidth---or, as we prefer to call it, his channel capacity (CC)[iii].

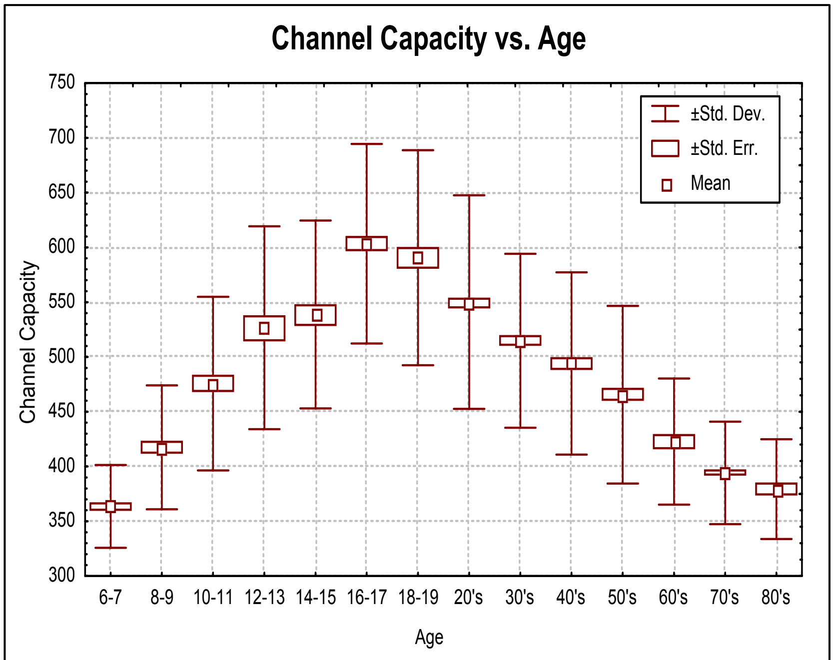


**Figure 3** CC increases from age 6-7, apparently geometrically, to a peak at age 16 –17, and then declines linearly into the 80's.

The driver---at the center of a feedback loop---adjusts his driving speed to a comfortable level with the goal of accommodating his own personal CC. The faster he drives, the faster the incoming information about the road, other vehicles and potential hazards---and the greater the CC he or she will need to safely process that information. No wonder, then, that teenagers, with peak CC on the WP test (see Figure 3), famously drive faster than any other group on the road[iv]. *They do it because they can do it.* Likewise, older people tend to drive more slowly, in keeping with the linear decline in CC as people age.

Despite the well-known correlation between age and crash frequency, WP does not use a driver's age as a predictor of safe driving. Age, a count of the earth's rotations about the sun (!) is a highly variable *measure of a group's capability,* whereas CC is a specific measure of an *individual's capability.*

As B. F. Skinner said in describing his approach to the science of behavior, "No one goes to the circus to see the average dog jump through a hoop. . . "[v] All along, WP has been developed using this "single subject" approach rather than averages among individuals.

THREE COMPONENTS OF SITUATIONAL AWARENESS.

Absolute speed on WP is necessary for predicting a driver's crash risk, but it is not sufficient. A second factor, "situational awareness" (SA), can mitigate or aggravate crash risk at a given driving speed. SA is defined here in terms of moment-to-moment *changes* in speed on the road or, in terms of the WP assessment, among the four screens. SA has three operationally defined components called impulsivity, sustained attention, and distractibility. All three predict crash proneness in a motor vehicle along with CC.

The three SA components are defined operationally in terms of response speed Some people are at their very fastest on the *first* screen of WP and then fall off sharply on the second screen. Colloquially, such individuals would be called "impulsive" or a case of "fools rush in where angels fear to tread". At the other end of the impulsivity dimension, where an individual is *slow* on the first screen and *much faster* on the second, one might say that he has a slow warm-up to novelty. At the extreme, "deer in the headlights" would be an apt description. Like CC, SA changes with age. Teenagers (16 to 20), a group well-known for their crash frequency, were found to be significantly more impulsive than drivers in their early twenties.

The second component of SA is sustained attention. Some individuals, for example, negotiate the second screen quickly and then fall off sharply on the third, making many errors. As with the impulsivity measure, teenagers are significantly less sustained than drivers in their early twenties.

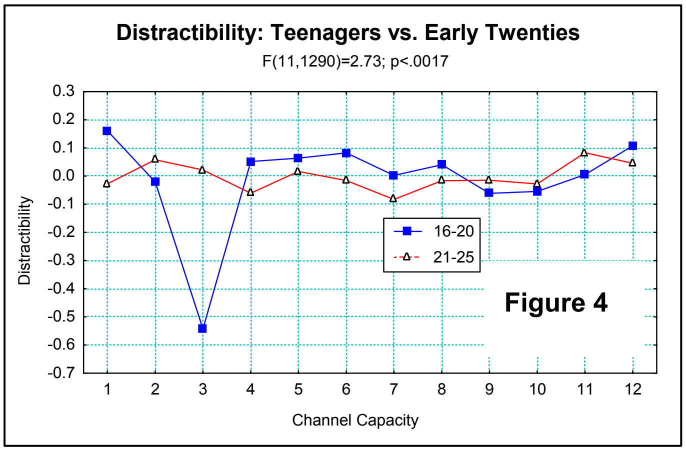


**Figure 4**

The third component of SA is distractibility, which, operationally defined, is the difference in speed between the third and fourth screens, where seven irrelevant pictures are interspersed among the numbers and letters. As with the other two components of SA, teenagers as a group are *significantly more distractible* than are young adults in their early twenties.

In summary, teenagers as a group have a dangerous mismatch between their CC and their SA. The average teen has a higher CC on WP than any other group *and drives faster*. At the same time, the average teen differs on the three components of SA compared to young adults *a mere five years older*. The discovery of this gap between CC and SA jibes with the drop in cost of auto insurance at age 25, and explains why teenagers represent only 7% of the driving population, while accounting for 14% of all driving fatalities and 20% of the injuries (NHTSA, 1996).

## INNOVATIONS

The ability of WP to accurately identify crash-prone drivers stems from test design innovations as well as theoretical and analytical developments which emerged under this Safety IDEA project .

## METHODOLOGICAL INNOVATION.

The basic trail-making task (Trails B in the Reitan test battery) has been known to correlate with driving skill at least since 1997[vi]. By presenting *four* screens of letters and numbers instead of just one by the Trails B, WP uniquely provides three new measures of SA along with a unique measure of CC. That WP requires but four (4) minutes from start to finish makes it one of the very briefest objective tests available.

### *THEORETICAL INNOVATION*

The power of WP to identify unsafe drivers stems from its theoretical underpinnings and the analysis that follows from it. In essence, the theory states that operators of motor vehicles are safe and competent to the extent that the moment-to-moment *demands* of the information streaming through the windshield are met by the moment-to-moment *ability of the driver to* process that information, *react to it in time, and thereby avoid a collision.*

Since the driver largely "chooses" the *demands* on him or her according to what he perceives as "safe and sensible driving", we can quantify his crash risk by precisely quantifying the information processing abilities on which his "safe and sensible" decision is made. If, for example, a driver habitually drives faster than his SA allows her to react, the research shows that crashes are more probable. This profile, as we have said, is over-represented among teenagers.

In validating WP, the task has been to quantify human information processing, an objective that experimental psychologists have pursued since the field was founded by Gustav Fechner nearly 150 years ago. My own basic research[iii] and the product developed under this contract has employed information theory as first described by the great applied mathematician, Claude Elwood Shannon. The application of information theory to the task of driving[vii] allows us to precisely quantify what has been said about CC and SA and to combine the two into a single risk metric.

***Flow Uncertainty.*** Imagine a world where all drivers have the same CC. In theory, they would drive in synchrony and never crash. Everyone would "go with the flow". In reality, CC is normally distributed among drivers and, according to our data, ranges from 0.5 bits to 5.5 bits/sec. Those near the mean CC among all drivers, even without correcting for driving exposure, crash significantly less than those at the extremes of CC (see, for example, Figure 2 b, Bin 7). So, a driver's own Flow Uncertainty is minimal to the extent that his or her CC leads him to synchronize with the average driver. Quantitatively, an individual's Flow Uncertainty is her CC normalized with respect to that of the driving population. A driver whose CC is at the exact median of the driving population has zero Flow Uncertainty. Flow Uncertainty *increases* with each "just noticeable difference" in CC in either direction from the *population* median.

***Driver's Uncertainty*.** In a world where everyone has the same CC and, in addition, unwavering SA, there would also be no collisions. No one would "leap before looking", nothing would distract from the task of driving, and no one's attention would flag out of boredom. Such drivers would have their vehicle under control from moment to moment. Driver's Uncertainty, as we call it, is the *change* in driving speed, i.e. the sum total of the three SA components, *relative to the individual's average driving speed*. Flow Uncertainty, by contrast, concerns the driver's typical speed *relative to other drivers nearby.*

**$H_{Total}$ = 1.0**

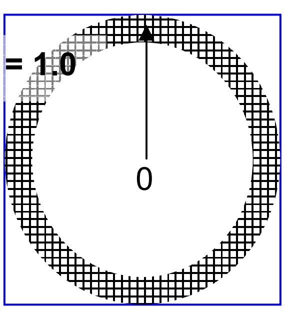


**Figure 5**

*The WP scoring algorithm adds Flow Uncertainty* ($H_{Flow}$) *and Driver's Uncertainty* ($H_{Driver}$)*, representing four different measures, to give* Total Uncertainty ($H_{Total}$). The individual driver crashes to the extent that his or her $H_{Total}$ approaches 1.0 (cross-hatched area). When $\mathbf{H_{Total}}$ is significantly less than 1.0, crash risk approaches the minimum.

$$H_{Total} = H_{Flow} + H_{Driver}$$

where $\mathbf{H_{Driver}} = \mathbf{\Sigma}(H_{Impulse}, H_{Sustain}, H_{Distract})$

## INVESTIGATION

### *DRIVER SUBJECTS*

The original plan under this project was to test commercial drivers "blind", i.e. without our knowing their crash history. Only then, according to the plan, would actual crash histories be compared to our risk assessment. The "blind" approach was especially appealing to the trucking companies since it allowed them to evaluate the accuracy of the test for possible adoption. Without a blind trial there was somewhat less interest in participating . Also, it wasn't easy to find drivers with a substantial crash history. First, preventable crashes are relatively rare events. Among auto drivers, for example, fully 50% never have a single preventable collision in a lifetime of driving. Over-the-road (OTR) semi-truck drivers follow a similar pattern despite their much higher exposure, perhaps 125,000 miles/year. . Furthermore, when commercial drivers accumulate collisions, they are routinely terminated, especially when collisions are severe, again reducing the population of driver/subjects. One truckload company known for its safe drivers provided us with more than 100 driver/subjects. Only two had more than two crashes during their tenure! In addition to these issues, some companies were reluctant to release crash histories and, of course, not all drivers who were asked to participate actually volunteered.

**Table 1 Drivers by vehicle type in this project**

| Vehicle Type | Drivers Tested | Included |
|---|---|---|
| OTR | 328 | 255 |
| Local truck | 641 | 522 |
| Commercial auto | 249 | 158 |
| ***Totals*** | ***1218*** | ***935*** |

A total of 1,218 commercial drivers from seven different companies completed WP as part of this contract (Table 1). Of these, 935 had sufficient on-the-road exposure and a company-provided crash history to be included in the data analysis.

VALIDATION SAMPLE

When the project began, WP data for drivers of trucks, cars, and transit buses were not aggregated out of concern that the *benefits* of additional subjects would be outweighed by the *costs* of added error variance such as the following:

- We were persuaded that the information processing demands of driving each vehicle type, e.g. autos vs. OTR trucks, were substantially different,
- Vehicle types differ greatly in their percentage preventable vs. non-preventable collisions. The great majority of transit bus crashes, for example, are classified as "non-preventable". For automobiles, it is just the opposite. Only later did we conclude that the *balance* of preventable vs. non-preventable collisions depends upon the speed demands of the vehicle and the number of stops it makes.
- Available criterion measures differed greatly among vehicle types. In this study, for example, crash data were only available during the driver's tenure with the company. For personal auto drivers, on the other hand, we were sometimes able to get lifetime crash data. In other studies, e.g. NHTSA studies of elderly drivers, a 3-year MVR was available along with a measure of crash severity.
- Finally, we were aware that WP scores, particularly CC, differed significantly according to the vehicle driven (see Figure x). Later, it became clear that CC is *not* the primary predictor of crashes, though it is positively correlated with driving

speed, type of vehicle driven, crash frequency and frequency of moving violations. Nor is SA the primary predictor of crashes.

The primary predictor of crashes, as we later concluded, is the *interaction* of CC and SA. As such, WP is not a "more is better" assessment, e.g. the higher the CC, the safer the driver. Instead, WP is better characterized as a "does it match" assessment. Does the driver's SA *match* his CC?

All along, readers can appreciate the difficulties that we faced as validation progressed: Given a demonstrably significant but too-noisy relationship between WP score and crash frequency, we constantly asked whether the "noise" in the data was due to the quality of the criterion measure, the way that WP's four independent variables were calculated, or due to some maddening interaction amongst all five variables..

As we collected more driver data, refined the CC and SA measures and, made progress with the information-theoretic scoring algorithm, it became clear that driver data could be aggregated independent of vehicle type as long as the individual driver's crash risk was corrected for his or her driving exposure. Empirical evidence in support of aggregation will be shown below.

For OTR truck drivers, a minimum of two years exposure was required for inclusion in this study *if he or she had no preventable collisions*; a one-year minimum was required if he or she had one or more preventable collisions. For local trucks, vans, and commercial autos, a driver without preventable collisions was admitted to the study if he or she had a minimum of *one year of* driving exposure.

For personal auto drivers, since crashes are relatively rare events, we used lifetime preventable crash frequency corrected for age. Any "noise" that this introduces, e.g. by including one's teenage crashes, is "randomized out" by the greater exposure that a lifetime frequency gives compared to, say, a 3 to 5 year record.

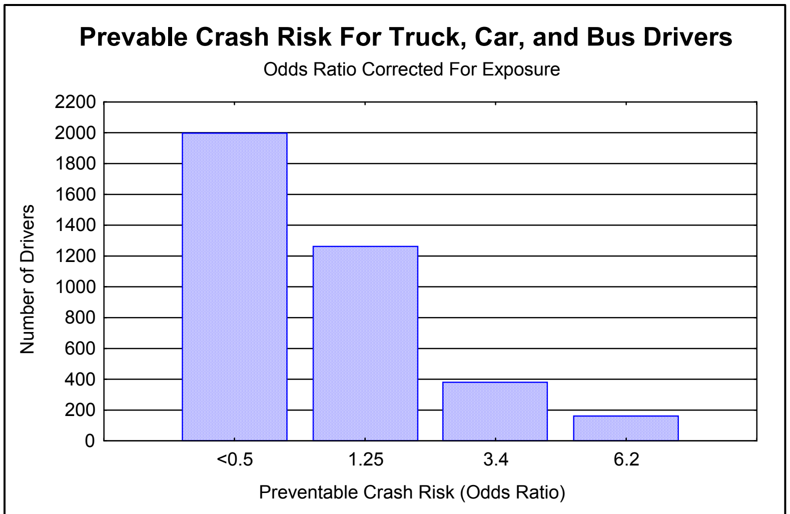


**Figure 6** Risk of a preventable crash is a Poisson distribution. Eighty-five percent of drivers are at or below the average risk of a preventable crash for all drivers.

The final criterion measure was an odds ratio, which was calculated by dividing a driver's exposure-corrected crash frequency by the average corrected crash frequency in the entire data base. So, a driver with an odds ratio of 1.0 is at average risk; 3.0, triple the average, and so on. *Four odds ratios* were discernable: a) 0.5, half of all drivers, b) 1.2, 35% of all drivers, c) 3.4, 11% of the drivers, and d) 6.2, 4% of the drivers. The Poisson distribution, where roughly 20% of drivers were responsible for 65% of the preventable collisions, is typical of crash data and other counts of rare events.

**Table 2 Percentage Vehicles in Full Data Base**

| Vehicle Type | Percentage |
|---|---|
| Auto - Personal | 51 |
| Auto - Commercial | 11 |
| Semi Truck - OTR | 13 |
| Semi Truck – Local Delivery | 2 |
| Van/Truck – Utility/Maintenance | 9 |
| Van/Truck – Refuse Pickup | 5 |
| Transit Bus | 8 |

Table 2 shows the percentage of vehicle types in the full validation data base of 3802 drivers. Each had a criterion measure of preventable crashes that was corrected for exposure.

**RESULTS**

Figure 8 shows mean CC for drivers of each of the vehicle types. Among the commercial truck drivers shown, note that OTR semi truck drivers, whose job demands fast driving, had significantly higher CC than local semi-truck drivers, whose job demands slower driving with many stops. Similarly, commercial pilots (n=550), not shown, whose job presumably demands higher information processing rates, had the highest CC of all the groups. Taken together, these findings indicate that, at a minimum, commercial motor vehicle operators in particular and commercial transportation operators in general are selected *and self-select* a job/vehicle according to CC---their own and that demanded by the job.

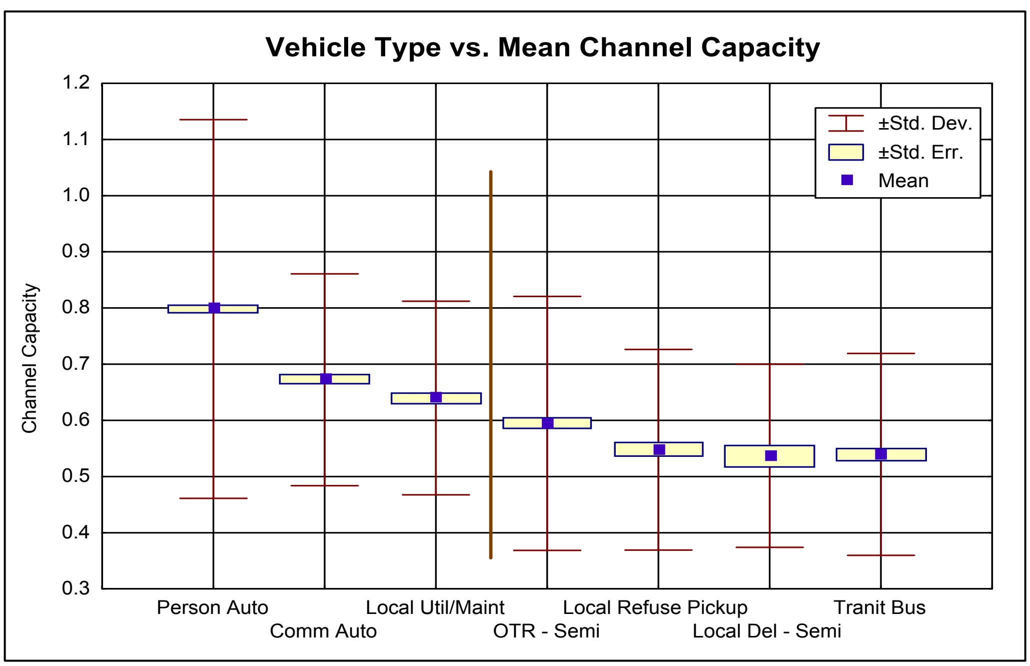


**Figure 7** Vehicle Type vs. Mean Channel Capacity.

Figure 9 shows that mean CC is not only specific to the various commercial vehicles, but also affects a driver's safety and therefore success at the job. Among OTR trucks, the *lowest percentage of drivers* with the very highest risk (filled squares) occurs in bins 3 to 6, near the group's mean CC (bin 4). Below that CC range (bin 2) the percentage drivers at highest risk spikes more than four fold. Above that CC range (bins 7 to 10), the percentage spikes up six-fold. Though CC bins 2 and 7-10 include only 24% of all OTR drivers in the data base, they account for more than 62% of the very highest risk drivers.

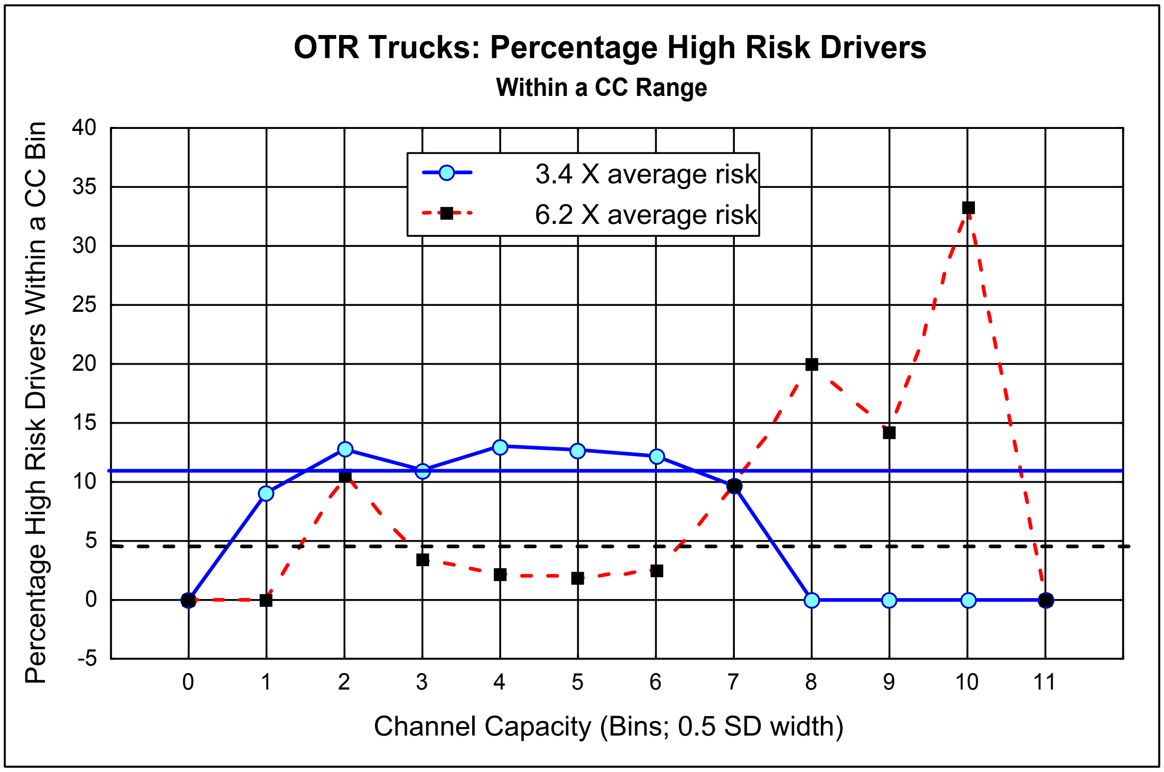


**Figure 8** The safest OTR drivers have a CC that falls within about ± 1 SD of the mean of all OTR drivers. Among the 24% of OTR drivers whose CC is above and below that range are 62% of the very highest risk drivers.

Importantly, this same U-shaped function was found *before* we added the 255 OTR drivers in this Safety IDEA project.

Among the pool of OTR drivers was a subset of 79 drivers who had tenure longer than five years; a count of complaints from the driving public and, importantly, a cost figure for each preventable collision. Though crash *frequency* was under-represented in bins 3 through 6, the highest *cost* crashes were over-represented in bin 4. Considering the 24/79 drivers whose CC fell in bin 4, 16% of them had crashes in the highest cost category, *more than triple the average base rate for all the bins. Indeed, they accounted for 57% of all the highest cost crashes, despite representing a mere 19% of all the drivers.* So, the middle of the CC distribution includes drivers with the lowest crash *frequency*, but among those low-frequency drivers is a cluster of drivers with the highest severity crashes.

Finally, the percentage OTR drivers with one or more complaints was below average among CC bin 4 and 5 drivers (again, the U-shaped function) and above the average in CC bins: 3 and 6 through 9. Complaints about bus drivers (n=299) follow the U-shaped function as well.

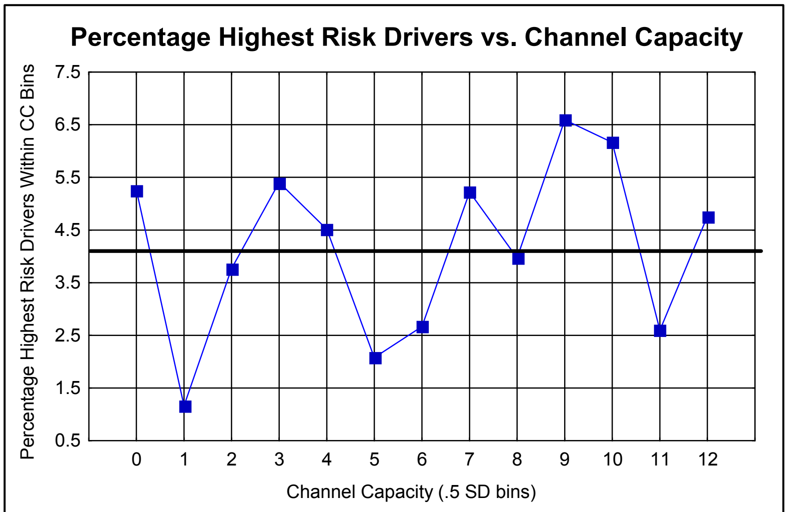


**Figure 9** Crash risk cycles as CC is incremented by a noticeable difference.

The function appears U-shaped for OTR drivers in part because the CC range is restricted. When we examine the distribution of high-risk drivers over the *full range of CC regardless of vehicle* type (51% personal auto), the function is noticeably zigzag or sinusoidal, as shown in the grouped data of Figure 10. Before showing the zigzag effect in the data of individual drivers, it is useful to consider why one would expect crash proneness to cycle as CC is incremented.

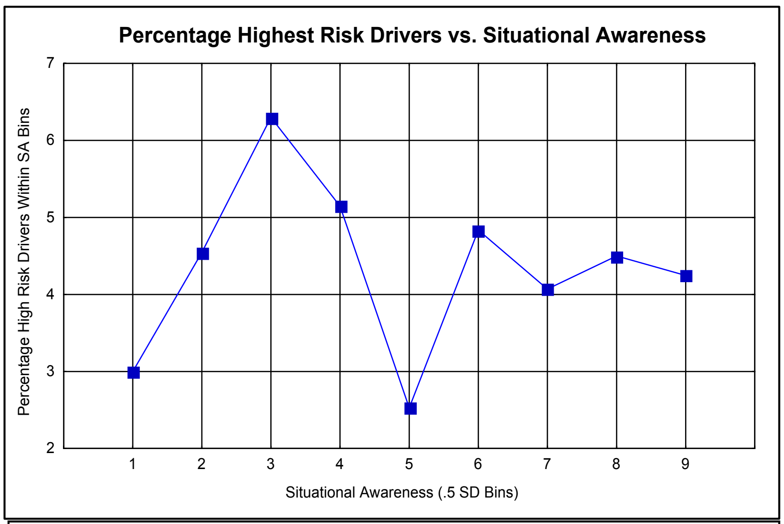


**Figure 10** Holding CC constant, crash risk cycles as SA is incremented by a noticeable difference.

A driver is at the center of a feedback loop that holds her collision risk at a comfortable level.[viii] She wants to get to her destination as soon as possible, but at the same time, she reins in her speed in service of speed limits, law enforcement and her perceived risk of crashing. There comes a point---a threshold---where the *benefits* of driving faster are outweighed by the associated *costs* of greater crash risk. And so she slows down. Among all drivers with the same CC, i.e. within a “just noticeable difference”, some will slow well *before* risk escalates and others, “after it’s too late” Drivers who routinely slow down too late are shown in Figure 8, and thus the cycling of the highest risk drivers throughout the range of CC.

The zig zag pattern is also apparent in Figure 11 where CC is held constant and SA is incremented by a “just noticeable difference”. Considering Figures 10 & 11 together, it doesn’t matter which component of uncertainty is incremented by a “just noticeable difference”, CC or SA. Both affect crash risk in the same way.

Could the results shown in Figures 10 & 11 be the basis for a predictive assessment of crash proneness? No. By incrementing both CC and SA at the same time---13 x 9 cells---the zigzags become sharper even as the number of drivers per cell declines. If we then identified which of the cells had the very highest over-representation of high risk drivers, we would find that those cells also include too many low risk drivers. The result would be an assessment with a very high---and unacceptable---false positive rate. As such, the analysis must focus on the data of individual drivers.

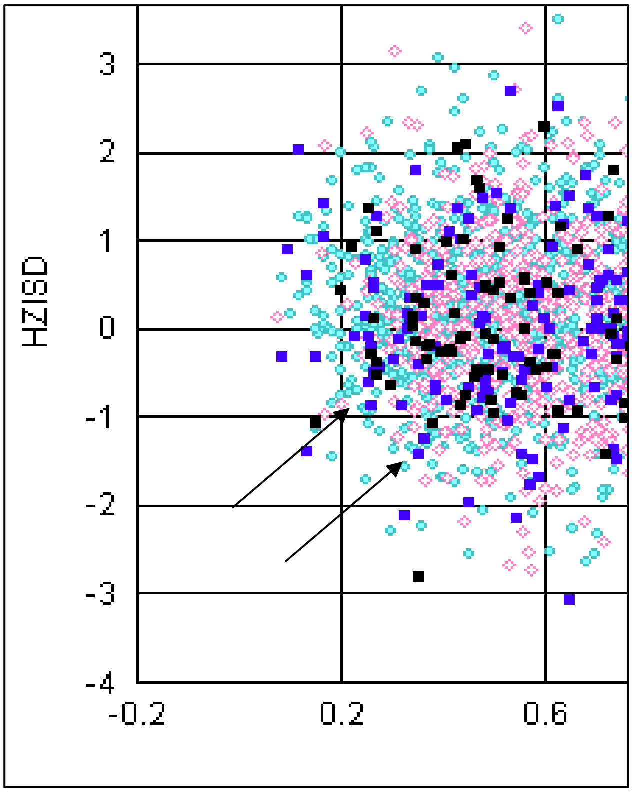


**Figure 11** Scatter plot shows individual drivers, regardless of vehicle type. CC, horizontal axis; SA, vertical axis.

Figure 12 shows crash data for individual drivers whose CC (horizontal axis) is below the mean of all drivers. The vertical axis is the aggregated SA measure. Black points are the highest crash risk drivers (odds ratio, 6.2); blue points, second highest (3.4 odds ratio); medium gray (1.2 odds ratio); and small gray (less than 0.5 odds ratio). Examples of the zigzag pattern are seen starting at the arrows.

Drivers whose combination of CC and SA falls along the “zig” tend to be high risk. Beyond the peak and along the same slope, drivers are low risk. The same slope, negative after the peak, is populated with high risk driver.

The zigzag shows that there is a critical *combination* of CC and SA that is associated with a high risk of crashing, though the combination is not peculiar to a given type of vehicle. Figure 12, from a region where all vehicle types are represented, shows a high-crash zigzag pattern; shapes and colors of points represent vehicle types as indicated. The mix of vehicle types in Figure 11 is approximately the same as that in the validation population.

**Table 3** Mix of Vehicle Types Forming Zigzag of Figure 12

| | |
|---|---|
| Bus | 6% |
| Car | 54% |
| Local | 18% |
| OTR | 22% |

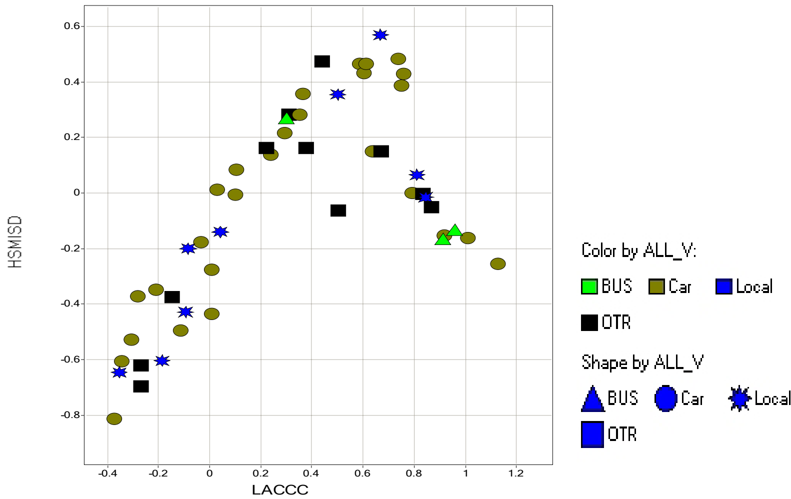


**Figure 12 CC vs. SA** The mix of vehicle types is approximately the same as that in the validation

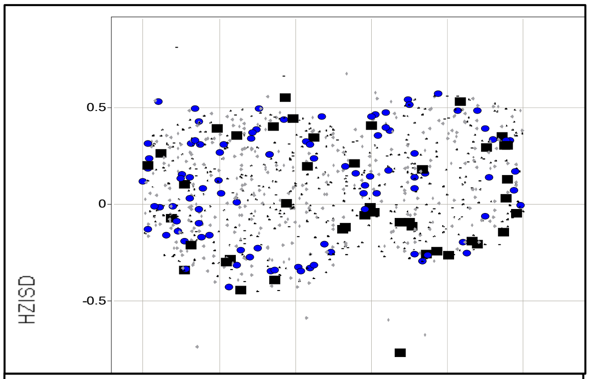


**Figure 13** Selected subset from Figure 10 showing that the zigzag can also be viewed as two contiguous ellipses or two sine waves 180 degrees out of phase

Figure 13, a subset of Figure 12, shows that what we have called "zigzag" can also be characterized as two sine waves 180 degrees out of phase.

Alternatively, the pattern can be seen as two ellipses. High crash drivers fall *on* the ellipse; low crash drivers fall *inside* the ellipse. From the standpoint of information theory, we can come to the substantial conclusion that a driver's preventable crash risk spikes up when his Total Uncertainty---the sum of his Flow Uncertainty (CC, horizontal axis) and his Driver's Uncertainty (SA, vertical axis), reaches a threshold defined by the perimeter of the ellipse.

So, at a given CC, a driver who "overdrives his SA"---like overdriving his headlights on a dark country road---has a greatly increased preventable crash risk. Colloquially, we say that he habitually drives "too close to the edge". *WP, for the first time, measures a driver's information processing with respect to just where that edge is located.*

A common method of validating assessment instruments is to develop the scoring algorithm on one group of subjects and then test it on a new group. Some of our validations have been conducted in that way. We have also used another strategy that exploits the zigzag or sinusoidal relationship that we have discovered.

The undulating pattern of high risk drivers is actually a Lissajous pattern. A Lissajous pattern occurs when sine waves, whether generated by an electronic circuit or by a violin, are mixed. In our data, the Lissajous pattern is generated by the addition of the four Gaussian crash predictor variables: CC, and the three components of SA, impulsivity, sustained attention, and distractibility.

Once the four variables are normalized, i.e. their weights are established, it can be shown that the left side of the scatter plot with respect to CC is isomorphic with the right side. Likewise, there is isomorphism along the SA axis. That one and the same functional relation, a sinusoid, prevails between the mix of predictor variables and the criterion measure is strong evidence that the test measures what it purports to measure, i.e. that it is valid.[1] Put another way, each cycle of the sine wave can be viewed as a replication of the "experiment".

Finally, the observed isomorphism supports the notion that WP is not a "more is better" assessment, e.g. faster is safer. More correctly, it is a "does it match?" test. Drivers can be safe at a wide range of CC (and presumably, driving speeds)---as long as their SA is a favorable match to their CC.

It has been suggested that WP could be "gamed" or "faked good" in the manner that personality or questionnaire-based assessments are "faked good". We don't believe that it would be easy to do, because the individual would have to know the exact relationship among the four variables and, more importantly, *make good* on what he knows by way of moving the mouse. If he can do that, he's a safe driver anyway. Could a driver candidate have a known safe driver "sit in" for her out of sight of the safety manager? Yes. However, the assessment takes but four minutes to administer and can be repeated (and witnessed) any number of times. A study addressed to the repeatability of WP using an earlier version of the scoring algorithm found that the "safe/unsafe" decision was, on average, consistent in 85% of the administrations.

In the course of validating WP, it has been useful to test operators of man/machine systems other than motor vehicle drivers. Our study of 80 jugglers bears special mention, since it is a useful metaphor for any activity where speed and accuracy are important. "Numbers jugglers" are performers whose goal is to catch and throw ever more objects. All jugglers, at the center of a feedback loop, constantly adjust the pattern of objects that they create. One would guess that the best jugglers are at the high end of CC, but in fact they are *not.* That is because speed generates error, i.e. variability in the pattern of objects. Variability is the juggler's enemy. In fact, the greatest numbers" jugglers---who can juggle 12 to 14 objects---are rock steady in their control of the throws and are not overly fast in their CC. They don't need to be fast (which just creates errors) to correct their errors, because, on average, they don't make errors. They occupy a sort of "sweet spot" where they "do it right the first time". Indeed, five Guinness Record holding jugglers all have precisely the same profile on WP. Plotting CC vs. SA, their data points fall on a 45 degree line ($R^2$=0.93).

The drivers whose data fall on the zigzag are analogous: their CC is *usually* just fast enough to compensate for the small errors inherent in their SA. However, given enough exposure, i.e. miles driven, their CC often falls short of matching their SA and they crash more as a result. The safest drivers habitually leave plenty of "room" for error.

---

[1] Were the opposite true, each individual would have a unique relationship between predictor and criterion measure, i.e. the pattern of high crash individuals would be random.

## PLANS FOR IMPLEMENTATION

The widespread offering of WP to the trucking industry was delayed as we perfected the scoring algorithm. Currently, the algorithm identifies 81% of the most crash-prone drivers (odds ratio, 6.2) and 69% of the next most crash-prone (odds ratio, 3.4). At the same time, the false positive rate, categorizing a driver as “high risk” (upper two odds ratios) when, *in fact*, he or she was “low risk” (lower two odds ratios) was only 8%. This level of sensitivity currently applies all drivers in the validation sample. As more validations are performed, accuracy will improve. There comes a point, of course, where one cannot be sure whether assessment “errors” are due to the insensitivity of the instrument or the quality of the criterion measure. Further validations will emphasize both.

We have contracted with an insurance carrier that will link their website to the WP assessment and test thousands of their customers, each with a known crash history. In the meantime, we will offer the fully functioning WP, first, at a discounted rate, to the companies that participated in this project and then to the rest of the trucking industry.

In parallel with testing drivers and developing the scoring algorithm under this project we have improved the on-line version of WP and included an on-line feedback report that has gone through several iterations. The report specifies the driver’s likely speed relative to other in his vicinity, her strengths and weaknesses regarding SA and, where appropriate, suggests specific actions that the driver can take so as to reduce the chance of a preventable collision. An abbreviated report is sent to management and includes the driver’s crash risk. A screen shot from an animated graphic for the automobile version of the report is shown below. A modified version for truck drivers will also be available.

**Figure 14** Animated windshield view; passing trees indicate speed. The user’s tendency to “go with the flow” based on his CC is indicated by a vehicle looming up ahead or in the rear view mirror. His SA is indicated by the size of an ellipse superimposed on the passing scene.

## CONCLUSIONS

We believe that WP will significantly reduce truck crashes in particular and improve safety in the transportation industry in general. In addition to the 1,218 commercial drivers assessed under this IDEA project, we have also assessed commercial and military pilots, pipeline controllers, nuclear plant operators, and airport security personnel, among others. The evidence indicates that the principles of information processing discovered under this project can be applied to operators of any man-machine system. WP will help organizations select safer drivers, better match an operator's skills to the demands of her job; and better identify fundamental features of an operator's performance that need remediation.

## INVESTIGATOR PROFILE

Michael B. Cantor, a human factors psychologist, is president of WayPoint Research, Inc. in Atlanta. He received a Ph. D. from the University of North Carolina at Chapel Hill in experimental psychology. Prior to his career in industrial psychology, which began in 1983, he was on the faculty and did basic research in information processing at Columbia University. Dr. Cantor received the Christopher Columbus Foundation Award in Homeland Security in 2003 for his work on identifying individuals with an aptitude for finding weapons on an X-ray screen.

WayPoint Research, Inc.
538 Burlington Road, Suite B
Atlanta, Georgia 30307
404 982 0011
404 982 9661 (FAX)
mbcantor@waypointresearch.com

REFERENCES

[i] Evans, Leonard *Traffic Safety,* Science Serving Society, Bloomfield Hills, Michigan, 2004.
[ii] Cantor, M. B. & Grasha, A  Unpublished technical report for the pharmaceutical industry, University of Cincinnati, 2000.
[iii] Cantor, M. B. Information theory: A solution to two big problems in the analysis of behavior. In P. Harzem and M. Zeiler's (Eds.), Advances in the analysis of behavior: Vol. 2. *Predictability, Contiguity and Contingency* (pp. 286-320), New York: Wiley, 1981.
[iv] National Highway Traffic Safety Administration, *Traffic Safety Facts*, 1996.
[v] B.F. Skinner,  *A case history in the scientific method*, 1956.
[vi] Stutts, J., Stewart, J .R. Martell, C. M.  Can screening for cognitive impairment at license renewal identify high-risk older drivers?  40th Annual Proceedings Association for the Advancement of Automotive Medicine, October 7-9, 1966, Vancouver, Canada.
[vii] Zeitlin, L. R.  A micromodel for the objective estimation of driver mental workload from task data.  Transportation Research Board, 1998 Annual Meeting.
[viii] Wilde, Gerald J. S, *Target Risk 2*, PDE Publications, 2001.