ABSTRACT

Driver distraction is a topic of considerable interest, with the public debate centering on the use of cell phones and texting while driving. However, the driver distraction/overload issue is really much larger. It concerns specific tasks such as entering destinations on navigation systems, retrieving songs on MP3 players, accessing web pages, checking stocks, editing spreadsheets, and performing other tasks on smartphones, as well as, more generally, using in-vehicle information systems.

Five major problems related to distraction/overload research and engineering and their solutions are addressed in this paper. Problems include (1) the misuse of the term distraction (and possible misdirection of effort), (2) driving performance measures and statistics that are either undefined or poorly defined (to be resolved by an SAE practice), (3) the workload of the driving task is not quantified (for which an equation is proposed), (4) the demand characteristics of in-vehicle tasks are not quantified (for which a scheme is proposed), and (5) too often, standards specify only measurement methods, not compliance criteria (which must be developed).

INTRODUCTION

Driver distraction/overload/workload is now the topic of considerable scientific and public debate. For the public, this debate has led to the passage of laws in cities and states restricting the use of cell phones and other devices while driving. Most recently, the province of Ontario, in Bill 118, took this step, becoming effective at the end of 2009 (Legislative Assembly of Ontario (2008). However, this topic has been one of scientific inquiry for some time (e.g., Brown, Tickner, and Simmonds, 1969).

Cell phones are but one example of nomadic devices that provide a wide range of features that are carried into motor vehicles. They supplement information and driver assistance systems now becoming commonplace. Each of these systems and their features is being provided for a valid reason, and collectively, they increase complexity of vehicle operation. With many more tasks for the driver to do, these systems could also collectively diminish usability and safety. This is particularly a concern for the carried in equipment, especially phones, which are rarely tested for use while driving, even though use while driving is expected.

Some perspective of how the current situation has arisen is useful. Consider the top-five selling vehicles in the U.S. in January 2010 per the Wall Street Journal (http://online.wsj.com/mdc/public/page/2_3022-autosales.html#autosalesC, retrieved February 25, 2010) (Table 1).

Table 1. Top Five Vehicles Sold in the U.S., January 2010

<table>
<thead>
<tr>
<th>Rank</th>
<th>Vehicle</th>
<th>Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ford F series</td>
<td>27,630</td>
</tr>
<tr>
<td>2</td>
<td>Chevrolet Silverado</td>
<td>22,772</td>
</tr>
<tr>
<td>3</td>
<td>Honda Accord</td>
<td>19,022</td>
</tr>
<tr>
<td>4</td>
<td>Nissan Altima</td>
<td>18,636</td>
</tr>
<tr>
<td>5</td>
<td>Toyota Corolla/Matrix</td>
<td>17,121</td>
</tr>
</tbody>
</table>

These vehicles are all reasonably well equipped. Among them, the 2010 Nissan Altima 2.5 with the SL package, a typical contemporary sedan, was selected as an example for which technical details are readily available. How does the current vehicle differ from the 2000 model? The 2.5S with the SL package and the technology option includes a navigation system with a 6.5-inch, touch-screen unit, traction control, and vehicle stability control. The navigation system
includes speed limit advisories, Zagat restaurant guide, XM traffic and weather data, a DVD player, and provides for several ways to accept auxiliary audio/video input. The Altima does not have adaptive cruise control, yet. Nothing like this was available on the 2000 model Altima, or for many other vehicles of that time period. The operation of these added systems is something drivers must now learn and will use while driving. If improperly designed, these systems can be distracting and their use overwhelming. Further, for any of the other vehicles in this list or a larger list, the same trend of increasing functionality should be apparent.

Each of the systems just mentioned, at least if it was developed by an automaker, was probably subjected to some sort of human factors analysis or evaluation. However, those doing the work are sometimes so close to those evaluations that they may not be aware of the larger problem of increasing total vehicle complexity. Moreover, because of the focus on very specific interface issues, engineers are not aware of the larger shortcomings of current work, which may fall below the quality of research and engineering being done in other industries. The author does not have any statistical evidence to support this assertion, though many human factors professionals will agree that the problems identified in the sections that follow are real.

Specifically, five major problems and their solutions are addressed in this paper. Those problems include (1) the misuse of the term distraction (and potential misdirection of effort), (2) driving performance measures and statistics that are either undefined or poorly defined (to be resolved by an SAE practice), (3) the workload of the driving task is not quantified (for which an equation is proposed), (4) the demand characteristics of in-vehicle tasks are not quantified (for which a scheme is proposed), and (5) too often, standards only specify measurement methods, not compliance criteria (that must be developed).

PROBLEM 1: MISUSE OF THE TERM DISTRACTION

Driver distraction and workload are often used interchangeably, but are not the same. Part of the problem is defining what is the problem. As has been stated before (Oberholtzer, Yee, Green, Eoh, Nguyen, and Schweitzer, 2007; Green, 2008), in the popular press but also in the scientific literature, the term “distraction” is often used to describe the topic addressed here. Two examples are the well-done series in the New York Times, “Driven to Distraction,” (http://topics.nytimes.com/top/news/technology/series/driven_to_distraction/index.html) retrieved March 2, 2010) and the most comprehensive book on the topic, Driver Distraction (Regan, Lee, and Young, 2009). Distraction generally refers to something that attracts and retains attention, whereas workload or overload refer to the individual and aggregate demands of the tasks a driver performs. In practice, sometimes the consequences of both are the same, but nonetheless distraction persists as the label for both phenomena, probably because it is easier to get attention and funding.

The naming/identification of the problem is important because of its implications for what one thinks the problem is and which performance measures should be collected. Keep in mind that there is just something compelling about answering a ringing phone, keeping a phone conversation going, responding to a text message, or completing an in-vehicle task such as entering a destination. When these tasks are conducted while driving, they become a safety issue.

PROBLEM 2: DRIVING PERFORMANCE MEASURES AND STATISTICS ARE EITHER UNDEFINED OR POORLY DEFINED.

Contemporary practice is to establish knowledge, such as how distraction/overload affects driving, based on the scientific method, a method that has existed for centuries. Richard P. Feynman, the Nobel prize-winning physicist at Caltech, in his well-known textbook Feynman Lectures on Physics (volume 1, page 2-1), said, “Observation, reason, and experimentation make up what we call the scientific method.” Following the scientific method involves creating a hypothesis to explain a phenomenon, collecting observable, quantifiable data in experiments to test the hypothesis, and using reasoning to interpret the results.

Those quantifiable data, measurements, must be repeatable and reliable. The lack of such measures has been a major problem for driving research and engineering, especially for work on distraction/overload. A few examples from Savino (2009), this master’s thesis, make the point. Savino reviewed the refereed human factors literature relating to driving, with the goals of determining the names used to identify common driving performance measures and statistics and how they were defined. He examined every issue of Human Factors and Ergonomics from 2000 to 2005, as well as the HFES and Driving Assessment Conference Proceedings, and other references. He did not examine SAE or ISO standards, as those interested in the research were familiar with their content. Overall, Savino examined 498 references, of which 111 were relevant to his research. Three examples of problem measures and statistics follow.

EXAMPLE 1: LANE DEPARTURES

The number of lane departures per unit of time or trial is a very common safety statistic. If drivers are distracted, they are more likely to depart the lane (Zhang and Smith, 2004). Table 2, from Savino’s research, shows how often various names for this statistic appear in the literature. Notice that the
most common name is used only one-third of the time, and it is not the preferred name.

Table 2. Names Used for Lane Departure in the Literature

<table>
<thead>
<tr>
<th>Measure name</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>lane deviation</td>
<td>5</td>
</tr>
<tr>
<td>out of lane</td>
<td>4</td>
</tr>
<tr>
<td>lane departure</td>
<td>3</td>
</tr>
<tr>
<td>lane exceedence</td>
<td>2</td>
</tr>
<tr>
<td>lane keeping</td>
<td>1</td>
</tr>
<tr>
<td>lane violations</td>
<td>1</td>
</tr>
<tr>
<td>crossing white sidelines</td>
<td>1</td>
</tr>
<tr>
<td>major lane deviations</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
</tr>
</tbody>
</table>

However, what is most critical is that this measure was rarely defined and when it was, the definition was imprecise. Contrast the definition of Jenness, Lattanzio, O'Toole, and Taylor, 2002 (page 594) of a lane departure beginning when “the automobile crossed the white sidelines on the roadway,” with that of Blanco, Hankey, and Chestnut, 2005 (page 1977) as beginning “when the vehicle's tire came into contact with the lane marker.”

So, what then is a lane departure? There are two aspects that need to be defined, what part of the lane is considered the boundary and what part of the vehicle is considered to have departed. In the U.S., as defined by the Manual of Uniform Traffic Control Devices (U.S. Department of Transportation, 2009), a lane marking is typically four inches wide. Depending upon whether the boundary is the inside edge of the line in the lane in question or the midline of the line is a two-inch difference. Further, the question is if departing is relative to any part of the vehicle, for example the outside edge of the mirror, or the outside of a front tire (Figure 1). For passenger cars and pick up trucks, the outside edge of the mirror is three to six inches farther outboard than the edge of the tire patch.

Thus, there are at least two candidate criteria for a lane departure, (1) the outer edge of the exterior mirror passes over the midline of the lane marking, and (2) the front tire touches the inside edge of the lane marking.

The first criterion is the most crash relevant. The second is easier to detect (when using a side-mounted camera). Simple math suggests there is a one to four inch difference between the two criteria. In reality, the difference is slightly less because most vehicles approach the lane marking at an angle. As a footnote, the definitions in SAE J2808 (Society of Automotive Engineers, 2007) and ISO 17631 (International Standards Organization, 2007) refer to the tire contacting the lane marking.

Example 2: Time-to-Line Crossing

Time-to-line crossing, a key safety measure, reflects the safety margin for lateral control. When drivers are distracted, the minimum time-to-line crossing over a time window decreases (Jamson, Westerman, Hockey, and Carsten, 2004). Fortunately, this is a term for which only that name is used. At a superficial level, it seems simple to define, basically how long it takes the vehicle to reach the lane boundary. Time-to-line crossing is defined in standards associated with lane departure warning systems, SAE J2808 (Society of Automotive Engineers, 2007) and ISO 17631 (International Standards Organization, 2007), but papers on lane departure rarely reference those standards or their definitions as well.

There are actually at least three different ways time-to-line crossing can be defined: (1) as lateral distance divided by lateral velocity, (2) as an expression that includes lateral acceleration, and (3) as the complete trigonometric solution that considers the radius of curvature of the vehicle's path and the radius of curvature of the road. Of the values provided by the three expressions, the first two of which are approximations, and all three can differ considerably. (See Godthelp, 1984; Van Winsum, Brookhuis, and de Waard, 2000).

Further, as with lane departure, there is the issue of what is considered a departure. For time-to-line crossing, it is typically when the tire touches the lane boundary, but it does not have to be.
EXAMPLE 3: HEADWAY

The more closely one follows a vehicle ahead, the more likely a crash (Ervin, Sayer, LeBlanc, Bogard, Mefford, Hagan, Bareket, and Winkler, 2005). Of the various types of crashes, rear-end collisions are much more likely when drivers are distracted (Wang, Knipling, Goodman, 1996).

In the civil engineering literature, for example, the Highway Capacity Manual (Transportation Research Board, 2010), headway refers to the time difference between when two successive vehicles pass by the same point on a road. Generally, it is the front bumper to front bumper difference (Figure 2). Civil engineers are interested in measuring traffic volume, how many vehicles/lane/hour occur on a road. The inverse of the interarrival time of successive vehicles is a measure of traffic volume.

![Figure 2. Gap vs. Headway](image)

However, human factors engineers are interested in the space between vehicles, the stopping distance available to avoid a crash. They often call that distance (or time) “headway,” as shown in Table 3, event though gap is the classically used and appropriate name (as in gap acceptance studies; Farber, Landis, and Silver, 1968; Chan, Ragland, Shladover, Misener, and Marco, 2005; Hwang and Park, 2005; Farah, Bekhor, Polus, Toledo, 2009). Strangely, SAE and ISO documents refer to gap as clearance, a name no one seems to use in the literature.

<table>
<thead>
<tr>
<th>Measure name</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>distance headway</td>
<td>6</td>
</tr>
<tr>
<td>following distance</td>
<td>3</td>
</tr>
<tr>
<td>space headway</td>
<td>1</td>
</tr>
<tr>
<td>vehicle headway</td>
<td>1</td>
</tr>
<tr>
<td>gap</td>
<td>1</td>
</tr>
<tr>
<td>gap distance</td>
<td>1</td>
</tr>
<tr>
<td>min, following distance</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
</tr>
</tbody>
</table>

However, the problem is not just that the name headway is inconsistently used, but the intended use is uncertain because it is not defined. Savino (2009) found 10 definitions for distance headway and 18 definitions for time headway in the literature. The definition of Strayer, Drews & Crouch, 2003, page 27, is typical. Following distance as “the distance between the pace car and the participant's car.” Not specifying the points on the vehicle leads to ambiguity. In a simulator, it is sometimes easiest to give the distance from center of gravity (CG) to center of gravity, since CG related measures are often computed. However, the difference between gap and headway, as defined here, depends on the length of the lead vehicle. The difference in length between car and a tractor-trailer is close to 40 ft (12 m), quite substantial. Further, it would be awkward to define several gap values, each for a different length lead vehicle or relative to a different part of the vehicle. (The CG-to-CG distance is not equal to either gap or headway, but some weighted combination that depends upon the length of the lead vehicle, an undesired complication.)

So, what is the solution to this problem? The first step was Savino's (2009) review that provided data to show the lack of consistency and provided a basis for defining terms. Using that information, a subcommittee of the Society of Automotive Engineers Safety and Human Factors Committee is developing a set of standard term names and operational definitions for driving performance measures and statistics (SAE Recommended Practice J2499). The terms initially being considered are: accelerator response time, accelerator to brake transition time, brake response time, steering wheel reversal, distance gap, time gap, headway time, headway distance, time to collision, lane departure, lane change, lateral lane position, and time to line crossing.

There are several key points to be made about these definitions and the process to develop them. First, the SAE subcommittee is being diligent to make sure the definitions are based on current practice and the research literature, as well as criteria for good definitions, not mere opinions. Second, to document how and why particular terms and specific definitions were selected, a rationale document is being written. Third, in several situations, such as for lane...
departure and time-to-line crossing, there is a need for multiple definitions to satisfy user needs, which will be provided. To avoid ambiguity, the intent is for authors in the future to identify which definition was used in each case (for example, as defined in SAE document such and such, definition A).

However, the most difficult part will be getting engineers and researchers to use those definitions. As was noted earlier, the author would assert that the frequency with which research in journal articles on lane departure to collision avoidance cite the relevant SAE and ISO standards is low, though the statistical analysis has yet to be done. This is a problem, and may be even one worthy of further discussion. A solution is for those who review manuscripts to be extremely diligent about commenting on or rejecting manuscripts that do not refer to the SAE practice.

This problem of undefined or inconsistently defined measures is quite serious, making automotive human factors engineering and research appear second-rate, and makes it difficult to consistently assess the effects of distraction. What would chemistry be like if there were ten different names for acidity, the pH value could be computed three different ways, and when used, the authors did not identify how the acidity/pH characteristic, or whatever they called it, was determined?

**PROBLEM 3: THE WORKLOAD OF THE DRIVING TASK IS NOT WELL QUANTIFIED.**

To date, the demand of the primary driving task in most studies is typically described in general terms, for example, as demanding, or in some studies, as low workload and high workload. Other times, it is measured, but no single or even small set of measures or statistics is consistently used in the majority of studies.

As an example, one of the author's studies (Tsimhoni, Green, and Watanbe, 2001) evaluated the effects of workload on Head Up Display (HUD) use in a driving simulator. Workload was manipulated by varying the radius of the curve driven, with the implication being that smaller radius curves represented a higher workload. However, there was no direct measurement of workload.

Keep in mind that what is low, moderate, or high workload is relative. In parts of the upper peninsula of Michigan, moderate traffic is when a driver sees another vehicle. In Tokyo, moderate traffic is when traffic is moving.

Consequently, until now, there has been no consistent basis for comparing, say a study on a four-lane road in the U.S. with a study on a two-lane road conducted in a driving simulator. There have been studies where researchers have attempted to measure primary task workload, using peripheral detection tasks, the standard deviation of lane position, etc. Again, no single measure is used consistently, and in some cases, such as for the peripheral detection task; there is no agreement on the procedure. Fortunately, ISO is developing a procedure for peripheral detection.

This lack of consistent and reliable measures to quantify test conditions also does not reflect favorably on automotive human factors work.

Workload depends primarily on road geometry, traffic, visibility, and the road surface condition, each of which can be quantified. For example, civil engineers describe traffic in terms of volume (vehicles/lane/hour) and the percentage of vehicles that are trucks. Commonly, traffic volume is described in terms of Level of Service, which maps traffic volume into letter grade categories of A through F, where A corresponds to excellent driving conditions and F to failing conditions.

As part of the SAVE-IT project, UMTRI developed a subjective rating method to quantify the demands of the driving task and equations to estimate workload from a description of the road geometry and traffic (Schweitzer and Green, 2007). (For other ideas, see Hulse, Dingus, Fischer, and Wierwille, 1989; Green, Lin, and Bagian, 1994; and Kondoh, Yamamura, Kitazaki, Kuge, and Boer, 2007).

In brief, researchers presented drivers with clips of road scenes that drivers rated relative to two anchor clips showing light and heavy traffic. Those anchor clips were assigned values of 2 and 6 respectively. Subjects then rated a large number of clips of expressways, urban, and rural roads, varying in terms of the lane in which the driver was positioned, the Level of Service, and other characteristics. It is important to note that the clips were taken from the Advanced Collision Avoidance System (ACAS) field operational test (Ervin, Sayer, LeBlanc, Bogard, Mefford, Hagan, Bareket, and Winkler, 2005), and associated with each clip were 400 engineering variables. Using regression analysis, equations were developed relating rated workload to driving statistics associated with each clip. One equation worthy of note is as follows:

\[
\text{workload} = 8.07 - 2.72(\log \text{MeanRange125}) + 0.48(\text{MeanTrafficCount}) + 2.17(\text{MeanAxFiltered}) - 0.34(\text{MinimumVpDot}(0 \text{ removed}))
\]
Where:

\[
\text{LogMeanRange}_{125} = \text{Logarithm mean distances (m) to the same-lane lead vehicles. If no lead vehicle, mean distance} = 125 \text{ m}
\]

\[
\text{MeanTrafficCount} = \text{Mean number of vehicles detected (15 deg FOV degree field of view)}
\]

\[
\text{MeanAxFiltered} = \text{Mean longitudinal acceleration (m/s2) of the subject vehicle}
\]

\[
\text{MinimumVpDot (0 removed)} = \text{Min acceleration of lead vehicle (m/s2), excluding the case of no lead vehicle}
\]

All statistics were computed over a moving 30-s time window.

This equation accounted for 85\% of the variance of the workload ratings, remarkably good.

Research is in progress to validate this equation and extend it to other driving situations. Certainly, over time, this and other equations will evolve, and the measures in them may change as research is conducted. However, what is notable about this equation is that all of the measures of interest can be readily collected and computed in most driving simulators. Further, most vehicles equipped with an adaptive cruise control system should be able to generate the data needed by the equation as well. This is important because in many cases computing workload will not require much additional effort.

The hope is that in the future, when someone does a study or performs an evaluation, they would report estimated workload (e.g., 5.5 on the UMTRI workload scale, identifying the equation used). Admittedly, just because two driving situations have similar workload estimates does not make them identical, but having some basis for comparison of studies is better than none at all.

The approach proposed here is a simplification of the demands of real driving. It is well known that human performance involves use of at least four types of resources, though many partition what people do more finely (Wickens and McCurley, 2008). Commonly, the literature refers to visual, auditory, cognitive, and psychomotor (VACP) demands, an idea popularized by McCracken and Aldrich (1984). VACP demands are discussed further later in this paper.

What is interesting about driving, is that most of the time, one could argue that the four resources occur in a particular combination. Driving is clearly visual, and without seeing the road, one cannot drive (Senders, Kristofferson, Levison, Dietrich, and Ward, 1967; Sivak, 1996). Further, visual demands are tightly coupled with cognitive demands to process the road scene, and motor output to steer the vehicle and control speed. Although one could create test conditions where they are decoupled, those combinations do not occur very often. Hence, for determining the primary demands of the driving task, measuring the aggregate demand or the primary visual demand should indicate the demand of driving. Thus, the simplification of only measuring overall demand is reasonable.

To date, these equations have only been examined in a single experiment (actually two by the time this paper appears) and they are far from perfect. However, they are the best estimates available and they are good. Quite frankly, it could be a decade or more before the research basis for these or alternative expressions is complete, and other five years or so before adoption occurs. When that finally occurs, how will the research conducted over the next 15 years be integrated into future work if there is no basis for comparison of the primary driving task?

**PROBLEM 4: THE DEMAND CHARACTERISTICS OF IN-VEHICLE TASKS IN QUESTION ARE NOT WELL QUANTIFIED.**

A topic of significant debate in the literature is what levels of task demands (especially visual, cognitive, and psychomotor) are excessive. However, because tasks are only described qualitatively, a quantitative answer is unlikely to appear. Further, when efforts are made to quantify the VACP demands of secondary tasks, they are often with reference to the interference of other secondary tasks (Wierwille and Gutmann, 1978; Casali and Wierwille, 1983; Zeitlin, 1995; Verwey, 2000; Harms and Patten, 2003; Jahn, Oehme, Krems, and Gelau, 2005). However, there are few tasks used as benchmarks consistently across experiments. The Alliance of Automobile Manufacturer guidelines uses manual radio tuning as a benchmark, but across studies, the total time for that task varies by a factor or six, hardly a stable value (Shah and Green, 2003).

There is reason to believe that in contrast to the primary driving tasks, the combination of VACP demands vary with the secondary task. Yee, Nguyen, Green, Oberholtzer, and Miller (2007) examined a large number of naturally occurring secondary driving tasks that drivers often do, such as smoking, eating, etc. These tasks had been identified in a prior evaluation of real-world driving (Oberholtzer, Yee, Green, Eoh, Nguyen, and Schweitzer, 2007). Tasks were split into subtasks (reach for cigarette, light cigarette, etc.), and their VACP demands were identified relative to UMTRI-enhanced scales that are part of the Army IMPRINT (Improved Performance Research Integration Tool, http://www.arl.army.mil/www/default.cfm?Action=445, retrieved February 28, 2010) benchmarks. IMPRINT is an application developed by the Army used to assess the workload of a wide variety of systems including future strike fighters (Brett, Doyal, Malek, Martin, Hoagland and Anesgart, 2002), future tanks (Mitchell, Samms, Henthorn, and Wojciechowski, 2002).
2003), and unmanned aerial vehicles (Walters, French, and Barnes, 2000) to determine manning requirements for systems and where interfaces and tasks need to be redesigned to optimize combat performance.

As part of the SAVE-IT project, the VACP scales were enhanced to make them easier to use. As an example, the visual scale is shown in Table 4.

Note that these scales are strictly ordinal, though it is common to use them as ratio scales. It may not be the case that to visually locate/align (selective orientation) is four times as demanding as visually register/detect image, only that it is more.

Furthermore, it is unknown how much demand is too much, either on a single scale or in combination. Under what circumstances is a visual demand of 6.0 excessive? That of course will depend on the workload of the primary task, the duration of the secondary tasks, and the cognitive, auditory, and psychomotor demands of the task as well. Of course, since these have yet to be quantified in a common manner in the automotive literature, there is no direct, quantitative answer to the excessive visual demand question just posed. However, should research begin to quantify tasks, then over time enough data will be accumulated to answer it.

As a first step to explore these scales, Yee, Nguyen, Green, Oberholtzer, and Miller (2007) examined the relationship between the VACP demands for a variety of naturally occurring tasks while driving. As an example, Figure 3 shows the relationship between subtask demands (again, lighting a cigarette was a subtask) for the visual and cognitive dimensions, the two most highly correlated. These correlations were computed assuming the scales were ratio scale, which they technically are not, but they commonly are used as ratio scales. In this analysis, subtasks were used because the combination of VACP values varied from subtask to subtask that formed a task. For example, the subtasks for a phone consist of preparing to use a phone to dial, answering a handheld phone, dialing a phone (case a-handheld, case b-hands-free), conversing on a phone, holding a phone but not conversing, and hanging up/ending a call.

### Table 4. Visual Demand Scale

<table>
<thead>
<tr>
<th>Rating</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>No visual activity</td>
<td>Self-explanatory</td>
</tr>
<tr>
<td>1.00</td>
<td>Visually register/detect image</td>
<td>Observe a warning light turn on</td>
</tr>
<tr>
<td>3.70</td>
<td>Visually discriminate (detect visual difference)</td>
<td>Determine which traffic light is on</td>
</tr>
<tr>
<td>4.00</td>
<td>Visually inspect/check (static inspection)</td>
<td>Check side mirror position while parked</td>
</tr>
<tr>
<td>5.00</td>
<td>Visually locate/align (selective orientation)</td>
<td>Change focus to a car</td>
</tr>
<tr>
<td>5.40</td>
<td>Visually track/follow (maintain orientation)</td>
<td>Watch a moving car</td>
</tr>
<tr>
<td>5.90</td>
<td>Visually read (symbol)</td>
<td>Read a native language</td>
</tr>
<tr>
<td>7.00</td>
<td>Visually scan/search/monitor (continuous)</td>
<td>Look through glove compartment</td>
</tr>
</tbody>
</table>

As a first step to explore these scales, Yee, Nguyen, Green, Oberholtzer, and Miller (2007) examined the relationship between the VACP demands for a variety of naturally occurring tasks while driving. As an example, Figure 3 shows the relationship between subtask demands (again, lighting a cigarette was a subtask) for the visual and cognitive dimensions, the two most highly correlated. These correlations were computed assuming the scales were ratio scale, which they technically are not, but they commonly are used as ratio scales. In this analysis, subtasks were used because the combination of VACP values varied from subtask to subtask that formed a task. For example, the subtasks for a phone consist of preparing to use a phone to dial, answering a handheld phone, dialing a phone (case a-handheld, case b-hands-free), conversing on a phone, holding a phone but not conversing, and hanging up/ending a call.

![Figure 3. Visual vs. Cognitive Demand for Various Subtasks](image)

**Figure 3. Visual vs. Cognitive Demand for Various Subtasks** Large dots signify multiple responses at the same data point.

### Problem 5: Too Often, Standards Only Specify Measurement Methods, Not Compliance Criteria.

Standards, guidelines, rules, and regulations fall into two categories, design oriented and performance oriented. Design oriented specifications identify specific physical characteristics for some feature and may specify values for it, such as a bumper height, a minimum acceptable contrast ratio for a letter, or a minimum intensity for a sound, say a warning. Performance specifications identify how well a system should do in a test, such as the maximum load on some body part in a crash, or the maximum allowable time for drivers to perform certain tasks while driving.

There are numerous standards, guidelines, rules and regulations that relate to driver distraction. The most important ones including U.S. Federal Motor Vehicle Safety Standards and NCAP regulations (U.S. Department of
Transportation, 2008a,b), national laws of countries other than the U.S., International Standards Organization standards, SAE standards and practices such as SAE J2364 (Green, 1999a,b; Society of Automotive Engineers, 2004), Alliance of Automobile Manufacturer Guidelines (Alliance of Automobile Manufacturers, 2003), state and provincial laws (Legislative Assembly of Ontario, 2008, California Department of Motor Vehicles, 2010; New York State Governor's Safety Commission, 2010), and local ordinances (City of Ann Arbor, 2010; http://www.clickondetroit.com/news/9231667/detail.html, retrieved March 2, 2010). This is a reflection of the complexity of producing a product that is manufactured internationally and sold in many jurisdictions, where many organizations have a rightful say in safety.

Which of these sets of documents is most important depends on this situation. However, it is apparent that ISO documents are consistently important, so they are discussed here.

Table 5 contains a list of the currently applicable ISO standards to driver distraction/driver workload. Of the 10 standards listed, four provide compliance criteria now and others may in the future. However, most of those criteria are very simple, easy to meet, and pertain to design-for character sizes, warning intensity, and so forth. Other than for the occlusion procedure (ISO 16673), what is excessively distracting, a performance characteristic, is left for the manufacturer or supplier to determine.

Not providing criteria, leaving up to the user to determine what is distracting, has some interesting consequences. The major automakers with a human factors staff have the capability to decide what is excessive, but where there is no performance criterion, there is no incentive to conduct these tests, so they may not do it. In the organizations with few or no human factors staff, they lack appropriate performance criteria, and accordingly will not perform the evaluation unless required to do so. To put it plainly, if there is no performance criterion for distraction testing, tests for distraction will not be conducted.

As a practical matter, automakers producing vehicles in the United States have agreed to comply with the AAM guidelines (Alliance of Automobile Manufacturers, 2003), and those in Japan comply with JAMA requirements (Japan Automobile Manufacturers Association, 2004; Nakamura, 2008). Aftermarket manufacturers should comply with SAE Recommended Practice 2364, but they do not, and there are consequences as a result (Canadian Broadcasting Company, 2010). Thus, when the critical testing occurs, it is primarily to meet the ISO 16673 requirements, similar requirements in the AAM guidelines, and the requirements of SAE J2364.

For those unfamiliar with them, the AAM guidelines for the design, installation, and use of telematics are based upon a set of 24 high-level principles (e.g., Principle 1.3 “No part of the physical system should obstruct any vehicle controls or displays required for the driving task.”). The AAM guidelines were derived from the 1999 European Statement of Principles, which have since been updated (Board and Stevens, 2001; Commission of the European Communities, 1999, 2006). The AAM guidelines go one step further by elaborating on their use and in some cases, providing test procedures. Of these principles, principle 2.1 matters most, (“Systems with visual displays should be designed such that the driver can complete the desired task with sequential glances that are brief enough not to adversely affect driving.”)

The AAM guidelines need to be improved. As noted by Morton and Angel (2005), page i “In the assessment of the AAM guidelines, principles appeared to be valid but often insufficiently detailed and too vague in the accompanying elaborations. This led to poor reliability of results between inspectors.” They further note on the same page, “With revisions the AAM guidelines may be sufficient to ensure safe operation of telematic systems, but insufficient in current form due to inadequate scientific support, incompleteness, and poor reliability of results.” Particular concerns were expressed with regards to the definition of measures and the selection of performance criteria.

So what should be done? ISO TC 22/SC 13/WG 8 needs to consider what it development program should be. If the purpose is to aid research by developing standard protocols, then that goal should be clearly stated. However, that will not occur if the committee consists largely of industry members, as it does now, with scant representation from the academic community. If the goal is to develop standards to aid industry and government in assessing the safety and usability of new products and services, then performance criteria must be included in standards.

Further, the AAM guidelines need revision, but that is unlikely at the moment given the current state of the automotive industry.

**SUMMARY/CONCLUSIONS AND CLOSING THOUGHTS**

This paper provides a rather critical perspective of some aspects of research and associated engineering pertaining to driver distraction/driver workload. It should not be taken to imply that those engaged in this research are not capable individuals nor that their intentions are not noble. In fact, many of the criticisms voiced apply to the author's own research. Certainly, the research to date has been informative.

Over time, research evolves, and the quality of work should improve as studies build upon the knowledge of prior activities. For that to occur, five things need to be done:
1. Authors need to clearly define what they mean by distraction to distinguish between the scientific and popular interpretations.

2. SAE needs to complete its work to define driving performance measures, and authors need to begin using them. Some means to induce use may be needed.
3. Studies of driving need to quantify the demands of the primary task of driving. The author has proposed one potential equation for that purpose.

4. The visual, auditory, cognitive, and psychomotor demands of secondary tasks need to be quantified. The author has proposed a method that uses the U.S. Army IMPRINT scales. The last three points have a common goal-to improve the rigor of evaluations conducted so they can be compared.

5. Compliance criteria need to be developed for several ISO standards and refined for the AAM guidelines. Without explicit criteria, everything is acceptable, and driving safety suffers.

REFERENCES


11. City of Ann Arbor (2010). *Use of Cell Phone or Other Portable Electronic Device While Operating Motor Vehicle or Bicycle Prohibited* (new section 10:1010a to be added to chapter 126 (Traffic of Title X of the Code of the City of Ann Arbor), version of February 28, 2010.


35. Legislative Assembly of Ontario (2008). An Act to amend the Highway Traffic Act to prohibit the use of devices with display screens and hand-held communication and entertainment devices and to amend the Public Vehicles Act with display screens and hand-held communication and entertainment devices and to amend the Highway Traffic Act to prohibit the use of devices with display screens and hand-held communication and entertainment devices and to amend the Public Vehicles Act with respect to car pool vehicles (Countering Distracted Driving and Promoting Green Transportation Act, 2008, Bill 118). Toronto, Ontario, Canada: Legislative Assembly of Ontario. (http://www.ontla.on.ca/bills/bills-files/39_parliament/Sessional/b118.pdf, retrieved February 27, 2010).


CONTACT INFORMATION

Paul Green
University of Michigan Transportation Research Institute
Driver Interface Group
2901 Baxter Road
Ann Arbor, Michigan USA 48109-2150
www.umich.edu
pggreen@umich.edu
1 734 763 3795