FINAL REPORT

A CRITICAL, COMPREHENSIVE REVIEW OF TWO STUDIES RECENTLY RELEASED BY THE OUTDOOR ADVERTISING ASSOCIATION OF AMERICA

Prepared for:

Maryland State Highway Administration
Under Project AX137A51

Through and with full concurrence of:

Positive Guidance Applications, Inc.
Gerson J. Alexander, President

By: Jerry Wachtel, CPE, President
The Veridian Group, Inc.
Berkeley, California

Consultant

October 18, 2007

A. BACKGROUND AND INTRODUCTION.

Outdoor advertising has been a fixture on America's roadways since there were roadways. From the "Burma Shave" signs and "Mail Pouch Tobacco" signs of the 1920s, to the jumbo neon spectaculars and computer controlled images on skyscraper curtain walls that grace Times Square, Hollywood and Las Vegas today, outdoor advertising has always been present in our built environment. Although the pros and cons of billboards have been debated for more than 50 years, it is perhaps their influence on traffic flow and safety that has been most controversial.

Now, a "perfect storm" is brewing that is bringing this issue to the forefront once again. Regulators and traffic safety experts are increasingly focused on driver distraction as a major factor in crash causation. Private roadway managers, together with traditional turnpike and toll road operators seeking new sources of revenue to support needed maintenance and safety improvements, are turning to advertisers to provide some of that revenue. And new technologies have brought large-scale, extremely bright, high resolution signs to the roadside – signs that can display a frequently changing series of images and are capable of presenting full motion video. Even more advanced technologies enable the advertiser to target roadside advertisements to the specific demographics of drivers passing these signs and to create signs that actually interact with drivers by presenting personalized messages on the billboard or contacting the driver's mobile phone.

Legal battles between advertisers and government agencies over the safety impacts of roadside billboards have taken place for more than two decades, and have reached the U.S. Supreme Court. After a hiatus of some years, these battles are looming again, but the stakes are now much higher because of the costs and revenues involved and the heightened concern about driver distraction.

B. THE PURPOSE OF THIS REVIEW.

In July 2007, the Outdoor Advertising Association of America (OAAA) announced on its website the issuance of two "ground-breaking studies" that addressed the human factors and driver performance issues associated with real-world digital (or electronic) billboards (EBBs), and the impact of such billboards on traffic accidents (Outdoor Advertising Association of America, 2007).

The OAAA website stated, in part, that this research "offers conclusive evidence that traffic accidents are no more likely to happen in the presence of digital billboards than in their absence." Since neither of these two studies had received public peer review at the time of their issuance, it was premature, at best, for the OAAA to make any claims of the validity of the findings. Also, since the accident study did not compare accidents in the presence of EBBs to accidents in their absence, the study presented no evidence, no less conclusive evidence, to justify the OAAA's claim. In addition, since even a cursory inspection of the human factors study showed that driver performance and behavior did, in fact, deteriorate when

EBBs were present vs. when they were absent, the OAAA claims seem difficult to support. Finally, the OAAA's claim that its Foundation for Outdoor Advertising Research and Education (FOARE) commissioned these studies to specifically examine whether there exists "a cause and effect link between outdoor digital billboards and driver behavior" demands scrutiny since the research methods and statistical analyses employed in these two studies were incapable of determining causality.

Nonetheless, news media throughout the country picked up on the OAAA web posting, and the debate over the pros and cons of these new billboards took center stage in many State and local government offices. In several cases, Government agencies at all levels came under immediate pressure to promulgate new regulations or amend existing codes and to permit the construction of new electronic billboards along their roads.

As a result of the issuance of these two studies and the claims made for them, and because of the need to address this technology by Government agencies nationwide, the Maryland State Highway Administration (MDSHA) asked this reviewer to perform an independent peer review of each of the two studies. This report represents the results of that review. The two studies are:

> "A Study of the Relationship between Digital Billboards and Traffic Safety in Cuyahoga County, Ohio," by: Albert Martin Tantala, Sr., and Michael Walter Tantala, Tantala Associates, Submitted to: The Foundation for Outdoor Advertising Research and Education, July 7, 2007", and
>
> "Driving Performance and Digital Billboards: Final Report," by: Suzanne E. Lee, Melinda J. McElheny and Ronald Gibbons, Virginia Tech Transportation Institute Center for Automotive Safety Research, Prepared for: Foundation for Outdoor Advertising Research and Education, March 22, 2007."

C. HOW THIS REVIEW WAS CONDUCTED.

Each of the two reviews that follows was conducted by thoroughly reading the report, performing independent verification, where possible, of the accuracy of statements made and data provided, reviewing the statistical analyses that were performed and the conclusions drawn from them, checking cited references and identifying important references omitted, and looking for signs of possible bias that might have influenced the decisions made or the conclusions reached. For both documents, our review, followed the same sequence: (1) Decisions and Assumptions Made in Support of the Research; (2) Methodology; (3) Review and Application of Cited Literature; (4) Statistical Methods, Controls, and Analyses; (5) Misleading and Inconsistent Reporting, and Evidence of Bias.

Because the strength of any research project rests heavily on the decisions and assumptions made in advance (*a priori)* by the researchers about what to study and what to exclude, and on the research methodology that rests, in large part, on those decisions and assumptions, we have given the greatest weight in our review to subsections (1) and (2) as listed above.

D. REVIEW OF THE TANTALA REPORT.

1. Decisions and Assumptions that Guided the Research

This study is based upon an after-the-fact (*post-hoc*) review of police traffic collision reports. The weakness of trying to understand issues of driver inattention or distraction (which is the real focus of this research) through a post-hoc review of summary crash data is discussed in detail in Subsection 2, below. In this Subsection, we address the decisions and assumptions made by the researchers that led to their determination of which crashes to include in the study, and which to exclude.

The decision of which crashes to study and which to ignore depends on two critical assumptions made by the researchers. The first assumption is based on their justifiable focus on those crashes that occurred in the vicinity of EBBs, i.e. within those roadway sections when an approaching driver could first *see*, and subsequently *read* the message on such billboards. In other words, the crashes of interest would be those that were initiated during the interval that an EBB was within the visible or legibility range of an approaching driver; and we would want to compare such crashes to those that occurred on comparable roadway sections where no EBBs were present. (As will be discussed below, however, such comparisons are not possible because of another assumption made by the researchers). Thus, the critical assumption underlying this decision is how to determine which crashes were, and which were not, within the visibility and legibility ranges of the seven (7) EBBs evaluated in this study. The second critical decision by the researchers that leads to the scope of included vs. excluded crashes for study is whether to include *all* crashes reported in these study and control sections of roadway, or whether to exclude some of the crashes for certain reasons. The researchers made the decision that certain types of crashes should be excluded from study because they were subject to certain types of biases. Below, we address each of these decisions, and the assumptions on which they were based.

a. Assumptions about the visual range and legibility range of EBBs

Let us first address the researchers' assumptions about the visibility and legibility distances of the seven (7) EBBs addressed in this study, which, in turn, led to their decisions about roadway sections for which to include and exclude crash data.

The authors use several different terms in their report to identify those roadway sections where approaching drivers might be able to initially see an EBB ahead and later read a message displayed on that EBB. Their terms include:

Two driving performance measures were examined for the nighttime data: standard deviation of speed and standard deviation of lane position. The standard deviation of speed appeared to be higher in the presence of both conventional and digital billboards than for baseline and comparison events. Lane keeping also showed a trend towards greater lane deviations in the presence of both digital billboards and conventional billboards.

The luminance values of many of the billboards, comparison events, and baseline events were also measured at nighttime. The digital billboards had noticeably higher luminance values than any of the other event types, even though their luminance was automatically reduced at night. This probably explains some of the driver performance findings in the presence of the digital billboards. The overall ranking of luminance by event (digital billboards were the highest, followed in order by comparison events, conventional billboards, and baseline events) closely mirrors the rankings of many of the performance measures for both daytime and nighttime, including eyeglance, speed maintenance, and lane keeping.

The overall conclusion, supported by both the eyeglance results and the questionnaire results, is that the digital billboards seem to attract more attention than the conventional billboards and baseline sites (as shown by a greater number of spontaneous comments regarding the digital billboards and by longer glances in the direction of the billboards). The comparison events, 25% of which included signs with digital components, showed very similar results to the digital billboards. Thus, there appears to be some aspect of the digital billboards and comparison events that holds the driver's attention, once the driver has glanced that way. This is most likely the result of the intrinsic lighting of these signs, which is noticeable even during the daytime. Drivers may also have maintained longer glances towards the digital billboards in the hopes of catching the next message (knowing that the message changes periodically). Although exploratory in nature, the nighttime results were very similar to the daytime results, with indications of degraded driving performance for digital billboards and comparison events.

These particular LED billboards were considered safety-neutral in their design and operation from a human factors perspective: they changed only once every eight seconds, they changed instantaneously with no special effects or video, they looked very much like conventional billboards, and their luminance was attenuated at night. It is thus quite likely that digital signs with video, movement, higher luminance, shorter on-message duration, longer transition times, and special effects would also be related to differences in driver behavior and performance. Because of the lack of crash causation data, no conclusions can be drawn regarding the ultimate safety of digital billboards. Although there are measurable changes in driver performance in the presence of digital billboards, in many cases these differences are on a par with those associated with everyday driving, such as the on-premises signs located at businesses. Conventional billboards were shown both in the current study and in the Charlotte study to be very similar to baseline and comparison events in terms of driver behavior and performance; thus, the design of digital billboards should be kept as similar as possible to conventional billboards.

"visible range from route," "viewer reaction zone," viewer reaction distance (VRD)," "viewer reaction distance zone", and "viewer reaction time (VRT)."

Unfortunately, the authors neither define these terms nor describe how the measures based on them were developed. The closest they come to providing this information is in their discussion of VRD, which they describe as follows: "... Viewer Reaction Distance (is) how far from a billboard that the driver is potentially within the "influence" of the billboard" (p. 45, and similarly on p. 79). In other words, viewer reaction distance is the distance in which the viewer can react. The authors further state that "reasonable values for VRD were previously determined in previous studies, and are a function of the driver's speed." This statement is incorrect. Viewer reaction *distance* is not affected by the driver's speed – the distance at which a driver can first see, and then read, any sign is independent of speed; it is only viewer reaction *time* that would be affected by speed. Further, the authors provide no citations for the previous studies mentioned, and offer no other basis for their development of their VRD formula. They report that, at 65 MPH, the VRD is approximately 0.2 miles with a VRT (Viewer Reaction Time) of 10 seconds (p. 79). A simple calculation demonstrates that, at 65 mph (95 feet-per-second or fps), 0.2 miles is traversed in 11 seconds, not 10. It is possible, therefore, that although this formula would yield an error approaching 10% (an assumed VRD of 10 seconds vs. a calculated VRD of 11 seconds at an assumed 65 mph speed), this may be the source of their assumption that VRD is 10 seconds. The merit of the researchers' values assigned to these critical measurements is further eroded by their failure to account for the fact that billboards on the opposite side of a multi-lane highway from an approaching driver provide a longer viewing time than those on the near side, and by their commingling of VRD with their measurement of "distance to the nearest billboard" (pp. 45-46) - a term which they do not define.

Surprisingly, the researchers' assumptions about VRD discussed above are directly contradicted by Table 2-3 on p. 15 of the report. This table is titled: "Visible Range of Billboards Along Interstate Routes;" it is never referred to in their report. The table shows the "visible range," in miles and feet, for each of the seven EBBs in the study. Although visible range is never defined, it would appear that these distances (which range from a low of 0.28 to a high of 2.15 miles) represent the maximum distance from which each of the EBBs can be seen by an approaching driver. Translating these distances to time, it can be seen that the billboard with the shortest visible range (#4) would be within an approaching driver's visual range for 15.6 seconds, whereas the billboard with the longest visible range (#5) would be visible to an approaching driver for 118.9 seconds, or nearly two minutes. (Note that both of these calculations are based on the researchers' claim that the Speed Limit was 65 mph). These data suggest that the researchers' decision to review only those crashes within 10 seconds upstream of any billboard is insufficient even to assess the potential influence of billboard #4, no less any of the other six, all of which were visible for greater distances.

The consequences for the validity of the data of this study are significant because the reviewers' assumptions led them to exclude all (an unknown number) crashes that were likely initiated in roadway segments further upstream from each of the billboards that they chose to study. For example, if the researchers determined that a particular EBB's zone of visual influence extends for 850 feet, they would exclude from their consideration any reported crash that took place outside this boundary. And the effect of these assumptions on the data considered by the researchers is substantial.

It should be obvious that every EBB visible along the route will have a different VRD and VRT depending upon numerous factors – sign size, height, and location, sight distance, size of characters in the display, roadway geometry, etc. Nonetheless the authors chose to assign each EBB the same VRD of 0.2 miles and the same VRT of 10 seconds. We believe that this is indefensible, especially given their data, in Table 2-3, showing the dramatically variable distances from which each of the seven EBBs can be seen. To take one example, if we assume (based on accepted industry practice) that 1" of letter height supports a legibility distance of 40', and a 14' tall billboard (all digital billboards in this study were 14' high x 48' long) with a character height of 75% of the available height or 10 feet 6 inches (a reasonable assumption based on the EBB images captured in Figures 2-4 and 2-8 of their report), then the legibility distance of such a sign would be 5040 feet, or 0.95 miles, nearly *five times* the VRD assumed by these authors. More significantly, the *visibility distance* is far greater than the *legibility distance,* and given the size, brightness, and frequently changing imagery on EBBs, it is reasonable to assume that crashes initiated within a given sign's visibility distance must also be considered. We have measured the visibility distance (day and night) to a recently erected EBB as greater than five (5) miles, and even though typical roadway sight distances would not permit such a lengthy visibility distance, it is reasonable to assume that the gaze of an approaching driver might be attracted to, and that such a driver might be capable of reading, an EBB at far greater distances, and for a far longer period of time, than the authors chose to evaluate in this study. We conclude, therefore, that the crash data accepted for inclusion in this study, based on the researchers' artificially constrained assumptions of VRD, has resulted in a substantial understatement of the true number of crashes that have occurred within the visibility and legibility range of the EBBs studied.

Because Viewer Reaction Zone is never satisfactorily defined, all of the results reported in Tables 4-1 to 4-4 must be considered suspect. Similarly, because the Visible Range is never satisfactorily defined, all of the results reported in Figures 4-2 to 4-9 must also be questioned.

b. Assumptions about bias in crash causes

The researchers made the assumption that certain classes or types of crashes should be excluded from study because they represent "bias". They use the terms "data-bias" and "interchange-bias." Data bias is discussed here; interchange bias is discussed below. Unfortunately, in 90 pages, there are no clear statements that identify exactly which types of biases were excluded, or why. Throughout the report, "examples" of such biases are mentioned in various contexts, but the list of excluded crash types changes from page to page, and a complete list is not provided. Thus, the reader is left to question whether other categories of crashes were also excluded from the data set. Excerpting from the examples given, the following crash types were excluded from the database:

- Deer hits (also called animal related) [1]
- Driving under the influence of drugs or alcohol[1]
- Adverse weather[2]
- Speeding[3]
- Senior related[3]

It is puzzling that the researchers chose to exclude crashes that occurred in adverse weather, and crashes that were "senior related." Since adverse weather is a known contributor to increased driver task demands, it represents the very kind of traffic/environmental condition for which one would want to study the potential distracting effect of such signs. Similarly, it has long been known that novice drivers and older drivers may have difficulties in highly demanding driving situations. For this reason, the decision to eliminate "senior related" crashes (no mention is made of crashes involving young drivers), raises additional concerns about the validity of the authors' approach.

c. Assumptions about interchange bias

The decision to eliminate crashes in the vicinity of interchanges (discussed by the authors on pages 49 and 77) is particularly troubling for several reasons. In their own words, they excluded interchange-related crashes because interchanges are "where drivers may undertake additional tasks, such as changing lanes, accelerating/decelerating, and negotiating directions." Since these demanding tasks associated with intersections are the very situations that are of concern to the traffic safety community, and since they are among the prime locations for high visibility billboards (because such locations may allow the billboards to be seen by traffic on multiple roadways), their *a prior* removal from this study is a cause for concern. It should be noted that the most recent billboard study published by FHWA (Farbry et al, 2001) and cited by these authors, specifically noted that intersections and interchanges were highly demanding road locations, and that such locations should be included in any study of electronic billboards.

---

[1] Discussed in Tables 4-5, 4-6, pp 45, 49, 77

[2] Discussed in Table 4-5, pp. 49, 77 ("snowfall" and "icy roads" on pp. 49, 77)

[3] Discussed in Table 4-6 (age 65 and above)

Decades of research into driver distraction demonstrates that alert experienced drivers can tolerate some distraction when their task demands are not high, but that all drivers have different cognitive capacities, and that there are certain road, traffic, and environmental conditions that may increase cognitive demands to the extent that additional sources of distraction should be avoided. Thus, the exclusion of some of the very types of crashes that might be *expected* to occur in the vicinity of EBBs is troubling, and, as with the decision to severely limit the distance upstream of EBBs in which crash data was collected, results in a likely substantial understatement of the actual crash statistics that took place in these roadway sections. Taken together, the choice of crash types to exclude is a serious weakness of this study, given that the very kinds of crashes excluded are those that would be of direct relevance to the potential for distraction caused by billboards.

Although the decision to exclude crashes in the vicinity of interchanges is problematic for the "temporal" study, it is far more harmful in that section of the report that deals with "spatial" factors. In Section 4.4, the researchers describe the "spatial analysis as having been performed to study "whether traffic accidents occur more frequently at or near digital billboards on specific routes." The same crash factor exclusions apply as elsewhere in the report, but here the reader is further misled because the researchers define the "exclusion zone" related to interchanges in two conflicting ways within the same sentence. They state that they excluded "those accidents and billboards on interchanges (entrances/exits) within one mile (1/4 mile on each side of an interchange)" (p. 78). Regardless of which distance was actually used, it is clear that any resulting findings are confounded by the fact that at least three of the billboards chosen for study (#3, Figure 2-8; #4, Figure 2-10; #7, Figure 2-16) are in close proximity to interchanges. Thus, if some percentage of accidents in the vicinity of these billboards was excluded due to the signs' proximity to the nearby interchanges, this artificially lowers the true number of crashes that may have been contributed by driver distraction due to these EBBs. As a result, the data for "bias adjusted" crashes in Tables 4-7 through 4-10, and in Figures 4-11 through 4-17 must also be considered suspect if not clearly insufficient.

Figure 1, (taken from the ClearChannel Outdoor website) shows the researchers' Billboard Number 3 and its proximity to an I-90 interchange. As can be seen, this sign is within ¼ mile of the interchange. So, regardless of what distance from an interchange was actually excluded in the data collection, this sign, and most likely, Billboards 4 and 7 as well, would have been excluded. Unfortunately, because the authors do not provide the reader with any information regarding the distance from the nearest interchange to any of their studied EBBs, there is no way to measure their proximity objectively.

Figure 1. Proximity of EBB #3 to an I-90 interchange. This image shows the same EBB depicted in Figure 2-8, p. 16, of the Tantala study. It is also Site # 22 from the Lee study. (Source: http://www.clearchanneloutdoor.com/products/digital/don/cleveland/index.htm)

In summary, the researchers' decision to exclude from study crashes that may have been affected by certain "biases" is essentially and critically flawed because it overlooks the most basic understanding of traffic crashes –that they are frequently multi-causal – and it is precisely when such multiple factors are at play – adverse weather, older drivers, complex interchanges, speeding) that the likelihood grows that the cognitive demands on the driver are increased and that irrelevant distraction cannot be tolerated. In other words, one should not exclude such factors because they cause "bias" – these are exactly the factors that interact to increase the likelihood of a crash when other factors such as inattention or distraction are present, and they must be investigated.

d. Assumptions about crash baseline data

Finally, one key assumption that colors the entire research paradigm is explained in two successive questions posed by the authors on p. 4 of the report. The first question asks: "...what is the statistical relationship between digital billboards and traffic safety?" This is a logical question to ask, and the one that should have guided this research. However, the next sentence, also posed in the form of a question, asks: "Are accidents more, less, or equally likely to occur near digital billboards compared to conventional billboards?" Unfortunately, it was this second question that guided the research, not the first. In other words, this study was not designed to investigate the safety of digital billboards compared to the *absence* of billboards; rather, it made the baseless assumption that conventional billboards were the acceptable baseline for comparison with EBBs. As a result of this assumption, the research methods failed to include true comparison sites where billboards were absent, and made any assessment of the contribution to crashes from EBBs against a true baseline impossible. Further, the announcement on the website of the Outdoor Advertising Association of America (OAAA) that accompanied the introduction of this report erroneously stated that this study "offers conclusive evidence that traffic accidents are no more likely to happen in the presence of digital billboards than in their absence." Clearly, since no comparisons were made between crashes in the presence and absence of EBBs were made, this claim is false,

2. Methodology.

This is a post-hoc accident study, meaning that the researchers reviewed summary reports of accidents that had been prepared by investigating police officers. Any such investigation is limited by the detail and accuracy of the reports reviewed, and by certain significant limitations in the accident reporting system itself, which the authors should have, but did not, acknowledge. For example, the majority of traffic collisions are never reported to, nor investigated by, the police, and thus any effort to determine a crash rate in the roadside areas studied will result in under-reporting. Second, and of critical importance for this study, unless a particular crash involves major property damage, serious injuries, or fatalities, the police investigation will most likely be rather cursory. In most States, a serious crash, and only a serious one, will require a specialized investigative team to examine the precursors to the accident (by reviewing evidence such as skid marks, debris fields, etc., and preparing a supplemental accident report); but for the vast majority of police investigated accidents, no in-depth investigation is performed. The result is that information on the traffic collision report form regarding the actual crash location is often incorrect, and leads to a serious weakness in the conduct of post-hoc accident investigations such as the present one. Specifically, police reports will almost always identify the location as the position where the involved vehicles came to rest after impact, since the costly and time-consuming in-depth investigations necessary to identify the originating location of the crash are simply not performed for the vast majority of crashes. Although the researchers are silent on this issue, it is almost certain that the crashes studied for this report are based on post-crash, at-rest vehicle positions rather than the upstream locations where the driver or drivers initially lost control or failed to pay attention. If our interest is in the "cause" of the crash, as it must be in a study such as this, the point of rest after collision is essentially meaningless – in other words, what we are concerned with when we study issues of distraction or inattention is not where the vehicles came to rest but where the crash sequence first began. As stated above, with the rare exception of serious crashes where supplemental investigations are performed in an effort to "reconstruct" the event, the researcher simply cannot know this information – and thus every identified crash location used for data analysis is suspect. In short, it is most likely that the traffic collisions evaluated in this study do not relate the triggers or precursors to billboard or EBB locations – they relate only the end result of such crashes to EBBs or billboards and thus they tell us nothing about the possible distraction effect of such roadside objects. This is enough to render these findings suspect.

The methodology included a "temporal analysis" which examined the incidence of crashes at locations where billboards had undergone conversion from traditional to digital display. Data was collected for 18 months prior to, and after, the conversion. Although this before and after analysis is necessary, it is not sufficient. Missing is any analysis of comparable sites in which there were either no billboards present, or billboards that were present but not converted. As has been shown in previous research in this field (e.g. Massachusetts Outdoor Advertising Board, 1976), it is possible that crash rates remained essentially the same in road sections featuring converted billboards, but actually decreased in sections that included non-converted billboards, or for non-billboard locations, during the same before-and-after study period.

Another key methodological weakness concerns the lack of definition and description of the assessment of "accident density" and the commingling of digital and conventional billboards in such an assessment.

3. Review and Application of Cited Literature.

The authors' include a list of 17 references, but none of these are actually cited in the text. Accordingly, the "References" section of the report is better identified as a bibliography. In addition, references made within the report of prior research are not accompanied by any citations, and thus it is not possible for the reader to know the source of the authors' claims.

4. Statistical Methods and Analyses Used

The researchers' discussion of crash statistics, including counts and rates, is misleading and erroneous. They rely on a review of police traffic collision reports but, as discussed above, fail to note that most accidents are not reported. On page 33 they define annual average daily traffic (AADT) as "the total volume of traffic in both directions of a highway or a road for one year divided by 365 days." Then on Page 47, in their discussion of Table 4-5, the authors discuss the number of collisions per year in the "digital-billboard locations" (which they do not define), and calculate the accident rate. However, because the EBBs studied for this report were single-sided (i.e. facing only one direction of travel), the authors have overstated the actual AADT by a factor of two, and the actual accident rate is therefore twice as high as reported.

In a section titled "Accident Density and Billboard Density," the researchers violate the basic tenet of their research objective by commingling digital with conventional billboards along the route. By including *all* billboards in their metric for billboard density, they invalidate any opportunity to compare, as their research objective says they will, digital billboards with conventional billboards, as well as any opportunity to compare digital billboards with the absence of billboards (which we believe to have been the proper comparison to have been made). Further, the authors' statement: "If a noticeable correlation between billboards and accidents exists, then one would expect a significantly larger number of accidents in areas with relatively high billboards densities" (p. 78). Aside from their misuse of the words "noticeable" and "significantly" in this context, this statement is incorrect because of the researchers' failure to control for the roadside environment (geometry, location of interchanges, presence of other roadside objects that might attract a driver's attention, etc.) in which billboards were present from areas where they were not, as well as their failure to separate digital from conventional billboards which might present quite different attention-getting characteristics.

The researchers misinterpret their own data. An examination of their Figure 3-5, for example, does not corroborate the authors' conclusions that "the median age of drivers involved in an accident are 23" (sic) or that "the winter months of 2005 had the most accidents on Interstates" (p. 33). Similarly, Figure 3-7 does not show that "the majority (of crashes) occurs during dawn and daylight hours" as stated on p. 34.

The authors misuse terms relating to their statistical analysis in an apparent effort to bolster their arguments. For example, they discuss a "noticeable correlation" – a term with no meaning. And, despite the fact that the authors know that correlation cannot imply causation, they nonetheless inappropriately suggest otherwise in several statements throughout the report (see, for example, pp. 2, 98).

Indeed, the researchers' entire discussion of correlation is apparently intended to suggest that no correlation less than 1.00 is indicative of any relationship. On page 81, for example, they state: "Statistically, a correlation coefficient of 0.7 or smaller is considered to indicate 'weak' correlations, at best, and does not indicate much difference from correlation coefficients of zero." On the contrary, any researcher who has studied traffic safety in a real world environment would be grateful to achieve results where correlation coefficients of 0.7 were found.

Because of their faulty assumptions discussed above, one of their major conclusions is simply unsupportable. They state: "if a noticeable correlation exists between billboards and accidents, then one would expect significant changes in the number of accidents between the 0 and 0.2 mile range and between the 0.2 and 0.4 mile range; the correlation coefficient would be large (close to +/- 1.00)." We take issue with this statement for two reasons. First, as discussed elsewhere in this review, because large, bright, colorful, changing EBBs can be seen, especially at night, at far greater distances than these researchers examined, and because it is likely to be these very sign characteristics that capture the attention of drivers long before a sign's message can be read, their potential to distract may occur far earlier (upstream) than the 0.2 mile cutoff used in this study. Second, it is generally understood that, in real world traffic safety studies, "strong" correlation coefficients rarely reach 0.6, no less approach 1.0.

The analysis performed in this study is based on what the authors call "commonly accepted scenarios relating accident density to billboard density, to 'viewer reaction distance,' and to billboard proximity (how far the accident is from the nearest billboard)." But none of these terms is defined, no references to prior research are provided, and the conceptual drawing used to explain these assumptions in Figure 4-1 (p. 46) provides nothing more than a visual illustration of this vague narrative. Thus, the reader simply cannot form an independent opinion of what was actually done, what assumptions were made, and how the data was collected.

5. Misleading and Inconsistent Reporting, and Evidence of Bias.

There are internal errors and inconsistencies throughout the report. Repesentative examples include:

- The key for Figure 2-1 is incorrect.
- Table 2-1 misstates the direction of the face of Billboard No. 3
- The description of Figure 2-2 (p. 8) is completely wrong
- The caption for Figure 2-2 does not depict what is in the Figure
- Figure 2-2 is identical to Figure 2-1 with key information missing
- All Latitude and Longitude data presented in the report is reversed
- The numbering system for the billboards studied is described as "arbitrary" (p. 10). This is not true. The numbering system (and all other descriptive information) is taken directly from the Clear Channel website.
- Data presented in the text disagrees with data in Figures and Tables.
- There are two tables each labeled Table 2-3 (pp. 11 and 15).

The authors base some of their calculations (e.g. "viewer reaction distance" and "viewer reaction time") on their statement that the posted Speed Limit was 65 MPH (p. 79). This is false. As clearly shown in their own Figure 2-10 (p. 17), the posted Speed Limit for at least one of their sites was 60 MPH. Although the authors provide no information about the posted speed limits at the other six sites, there is no reason for the reader to believe that they are other than 60 MPH. Of course, at 60 MPH, a driver approaching an EBB will be able to see and read the billboard for a longer period of time than would be the case at 65 MPH, thus further challenging the researchers' choice of an upstream cut-off point for crash statistics. Figure 2, taken from the ClearChannel Outdoor website, depicts the same EBB (No. 4) shown in the Figure 2-10 of the study. In addition to showing the posted speed limit, this figure clearly shows another billboard within the approaching driver's field of view, and the proximity of this EBB to the I-77 interchange ahead.

Figure 2. Image showing EBB #4 adjacent to posted Speed Limit signs. This image shows the same EBB depicted in Figure 2-10, p. 17 of the Tantala study. I-77 interchange signs can clearly be seen, as can an additional billboard in the driver's view. This is the same sign represented as Site No. 42 in the Lee report. (Source: http://www.clearchanneloutdoor.com/products/digital/don/cleveland/index.htm)

The authors present considerable data that are completely irrelevant to the study, while ignoring data of central importance. They take five pages (23-27) to provide superfluous information about the Interstate Highway System, provide data in Tables 3-1 and 3-2 that has no relevance for the study, and, on pages 29-31, repeat, verbatim, information previously provided on pp 10-11. The irony of this presentation is that the report is completely lacking in vital information about these EBBs and the studied road sections. For example, the authors provide no information about horizontal and vertical curvature, merges or lane drops, presence of official signage, or intersection characteristics such as entrances, exits, gores, etc. They don't even tell us how close any of the studied billboards are to the nearest interchange, despite the fact that their figures clearly show such proximity in at least three of the seven studied locations.

A number of the Figures and Tables in the report are taken directly from the website of ClearChannel Outdoor, the owner of the seven billboards studied. As indicated above, despite certain statements made by the authors, such as: "The numbering of the digital billboards in this study are (sic) arbitrary" (p. 10) this numbering system and other information relevant to each billboard is actually that provided by the billboards' owner, ClearChannel Outdoor, LLC. Further, in an article reported in the *Philadelphia Inquirer* (Slobodzian, 2007), a ClearChannel Outdoor executive stated that his company "hired" the researchers to perform this study in Cleveland.

E. REVIEW OF VTTI REPORT.

1. Decisions and Assumptions that Guided the Research

a. The decision to collect data for an 8-second interval prior to passing a site.

The researchers chose a time period of 8-seconds in advance (upstream) of a billboard, and, indeed, all sites, to record driver performance and eye glances. This interval was chosen, according to the authors, because this was the message-change cycle for the digital (electronic) billboards studied at these sites. The assumption that 8 seconds was a reasonable data capture interval raises several concerns, discussed below under Methodology.

b. The choice of "control" sites

The researchers selected four types of "events" or "sites" at which to collect data. For the main (daytime) portion of this study, there were 5 digital (electronic) billboard locations, which we have called "study sites," and three other types of locations, which we have called "control sites." These included conventional billboards (N=15), baseline sites (N=12), and comparison sites (N=12). The study authors provide no images or schematics of any of the 44 locations, and their descriptions and definitions of the site characteristics, particularly for the baseline and comparison sites, are vague and inconsistent, making it difficult for the reader to determine just how these site types compared. Baseline sites, according to a statement in the study Abstract, contained "no signs" (p. 6). But we later learn that this was not the case. Some baseline sites (the authors never state how many) in fact, contained signs. But the most serious problem is with the assumptions made for the comparison sites. On page 8 we are told that comparison sites are "similar to items you might encounter in everyday driving." On page 21, they are described as "areas with visual elements other than billboards." Later on the same page we are told that some of these sites included on-premise signs, variable message signs, and "digital components." Then, in Table 2 (p. 22) the reader can see that one comparison site is described as a "tri-vision billboard" and three others as an "on premise LED billboard." In reality, then, the comparison sites may have been just as visually compelling, if not more so, than the EBB sites that were the principle focus of this study. This intentional confounding of study and control sites seems designed to further the researchers' purpose of diminishing any adverse findings from EBBs by showing them to be no worse than existing sources of distraction.

As expected, the report's findings tend to bear out this concern in that, for many measures, EBB and comparison sites elicited similar results, and these results differed, often significantly, from those obtained at conventional or baseline sites. The problem for the researchers is how to treat these findings given their *a priori* site selection decisions; the problem for the reader is how to interpret them. In our opinion the approach adopted by the researchers is seriously flawed. It takes the clear evidence found in this study that digital billboards are associated with adverse driver performance, and twists these conclusions in such a way to suggest that there is no problem with digital billboards because drivers are already distracted by other "comparison" sites. The net effect of this is to avoid directly addressing the research objective, which was: "to assess the effects, if any, of digital billboards on driver behavior and performance" (p. 8), and instead answer a question that was not asked – how does driver behavior and performance compare for digital billboards vs. similarly distracting on-premise signs? In short, the authors' assumption that their carefully chosen comparison sites were appropriate control locations against which to compare the effects of EBBs enables them to slant their findings to suggest that, because driver performance in the presence of digital billboards is similar to their performance in the presence of these equally distracting "comparison" sites, there is no cause for concern regarding the safety of EBBs. Looking at the data, this seems to be a serious error.

c. The decision to minimize nighttime data collection.

Digital billboards are of particular concern to traffic safety experts at night, due to their ability to achieve high brightness and contrast levels, their high resolution imagery, and their visually compelling message changes, all of which can act to capture the attention of the driver at the expense of other targets in the scene (such as official signs and signals, pavement markings, and other vehicles). Thus, we question why the researchers chose to perform only a limited night-time study, one which included, *by design*, too few participants to enable them to analyze the data statistically. This decision is particularly troubling because, as might have been hypothesized, the researchers found indications of greater distraction by digital billboards vs. control sites at night with all of their measures, and, they suggest, at least some of these findings "would show statistical significance" in a larger study (p. 64). They state: "The overall conclusion (of the nighttime data), supported by both the eyeglance results and the questionnaire results, is that the digital billboards seem to attract more attention than the conventional billboards and baseline sites (as shown by a greater number of spontaneous comments regarding the digital billboard and by longer glances in the direction of these billboards)" (p. 10). They also report that both performance measures, speed maintenance and lane keeping, were poorer at night for digital billboards (and conventional billboards) than for comparison or baseline sites.

2. Methodology

a. Lack of control locations.

The researchers selected some study sites on the right side of the road and some on the left, then recorded and analyzed whether drivers glanced in the direction of these sites as they approached and passed them. In some cases they found examples of participants looking in the direction opposite to the site being studied. When such behavior occurred in the presence of billboard sites, they interpreted this to mean that the billboard was not attention-getting. But there is no evidence to suggest that they sought to identify or control for the possible presence of billboards or other visually attention-getting targets that may have existed along the roadside opposite their study sites. In other words, when they selected a study site on the right, they provide no information to suggest that they made sure that there was nothing on the left that might capture the driver's attention. If, in fact, they did not identify and control for such opposing sites, , then the eye glance data that they captured are suspect.. Since they do not report any efforts to evaluate and control for such conditions, one must assume that they did not do so. In short, it is entirely possible that glances to the left when a billboard was on the right (or conversely) were made because there was a competing, perhaps compelling, site across the road from the study site that was neither controlled nor evaluated.. A similar concern exists for uncontrolled sites that might exist on the same side of the road as a site of interest, and within a driver's field of view as he or she approached that site. Given the coarseness of the eye gaze data, there was likely no way for the researchers to know whether a particular participant was looking at the study site or an unidentified site for which they did not control. As one example, a review of Figure 1 shows the EBB of interest on the right side of the road, but the figure also shows a large billboard on the left side of the road that appears in the center of the image. If the researchers captured eye glances straight ahead or to the left at this location, they might have been due to the participant looking at this uncontrolled billboard.

b. Lack of controls within and across study sites.

Although the five EBBs studied each measured 14' high by 48' wide, we are given little information about other important characteristics of these signs; characteristics that could have had a direct impact on their attention-getting qualities, such as their height, angle to the drivers' line of sight, and proximity to the road. Further the reader is told little about roadway geometry, prevailing traffic speeds and volume, etc. Any of these factors may have affected the comparability of sites. Even though all five EBBs were 14' high and 48' wide (although mounted at very different heights relative to the road surface elevation) there was no consistency of sizing of conventional billboards or signs on the comparison sites. Indeed, we are told on page 21 that conventional billboards included a "few" that were of other sizes, including "standard poster, junior paint, and 10'6" x 36' bulletins." Since the size of a billboard or other sign likely has a direct relationship to the distance from which it can be seen and the size of lettering that it can accommodate, this failure to control for sign size and other characteristics relative to a sign's visible and legibility range is an important oversight. In our opinion, without any effort to control these basic site and sign characteristics, it is difficult for the researchers to defend any interpretations they may have made from their data.

c. The implications of the 8-second data recording period.

As discussed above, the authors describe four different types of "events" to which they studied driver response. Driver response to each event was recorded for a period of 8 seconds, ending at the time that the instrumented vehicle passed the event, and beginning 8 seconds upstream of that location. However, there are several methodological problems both with their choice of an 8-second interval, and with their ability to define and measure it.

At 65 MPH, the presumed speed on the freeways studied, a vehicle travels approximately 95 feet per second. Thus, during an 8-seoncd interval, a vehicle will travel 760 feet. The accepted practice for highway signs is that 1" of letter height can be read from approximately 40 feet away. So, for a billboard with 24" letters, the sign can be read from approximately 960 feet. In other words, a billboard can be *read* by the average driver at distances greater than those covered by the 8-second interval. When we recognize that letters on such signs may be considerably larger than this, and that such signs may display images of products or services dramatically larger than the size of letters (the signs studies are typically 14 feet tall), it should be obvious that these messages can be read at distances far greater than the authors' 8-second interval could accommodate. (Applying a scale to the word "Digital" in the signs shown in Figures 1 and 2 shows that these letters are 48" in height, indicating a legibility distance of 1920 feet. It takes more than 20 seconds to traverse this distance at 65 mph). In addition, because of the brightness, contrast, and image quality of digital billboards, and the fact that (in Cleveland) their message changes every 8-seocnds, it should be obvious that driver attention to the billboard may be initially attracted at far greater distances than those at which the message can actually be read. As a result, the choice of an 8-second data recording interval is likely to result in a substantial underestimation of the distracting effects of digital billboards compared to other roadside sites including more traditional billboards and on-premise signs.

The authors state that they chose an 8-second data collection period because the "digital billboards were programmed to change messages instantaneously once every 8 seconds; an event length of 8 seconds thus made it highly likely that a message change would be captured during the event" (p. 21). This argument is flawed for several reasons. First, as described immediately above, the sight distance and legibility distance, coupled with the size of the signs studied and their letter height, strongly suggests that digital billboards can be seen and read far earlier than 8 seconds upstream of the sign, thus strongly suggesting that the data recording interval should have been much longer. Second, it is obvious that had the researchers selected any data recording interval longer than 8 seconds, it, too, would have permitted them to capture a message change during each driver's approach to the event. Finally, despite their implicit recognition of the possible significance of a driver actually observing a message change during his or her approach to the EBB, the researchers apparently never actually recorded any data on message change, and therefore had no way to evaluate any possible driver response to it.

The seeming precision of the 8-second recording interval is also belied by the imprecision of the site characteristics. If a site contains a single billboard, one can measure the point in time when a driver passes that billboard. But how does one identify and measure the point of passage for the other types of sites studied? Since many of the comparison sites contained multiple objects, including multiple signs, there is no obvious end point. Similarly, for baseline sites which were devoid of any signs, one must question how an end point was defined.

Because of what is often called the driver's cone of forward vision, signs on the left (billboards or otherwise) can generally be seen for a longer time and distance than can signs on the right. Given that the same 8-second interval was applied to signs on both sides, this raises questions for the viability of the start time for data collection at each site, given that this start time was defined as 8-seconds prior to the end point.

Some signs are located perpendicular to the driver's direction of travel. Others, such as some two-sided billboards and many on-premise signs, may be located at other angles, including parallel to the driver's direction of travel (such as when mounted on a building façade). In addition, the lateral distance of each sign from the driver's line of sight varies greatly as a result of factors such as: lateral distance from the road edge, and the number and width of lanes, medians, and shoulders. If the same 8-second point for passing a sign was applied regardless of sign angle to and lateral distance from the road, then some signs would be visible to drivers for less time than others, thus rendering the 8-second recording interval inconsistent across the studied sites.

In summary, the researchers' choice of an 8-second data recording interval was inappropriate for many reasons, and resulted in unequal exposure to signs of interest across sites. A more appropriate way to determine the data collection interval would have been to identify the point at which a billboard or other sign of interest fell outside a predetermined angle of view from the driver's line of sight along the road axis, and defined the recording interval upstream from that point. This would have assured a more equitable, and comparable, identification of sight distance and would not have had the effect of artificially reducing the available glance times and control measurements made for the signs of interest in this study. Of course, it also would have made data actual collection more challenging.

d. Measurement of nighttime luminance levels

The authors measured the luminance levels of different sites at night. They took these measurements from the participant-driver's eye position, a decision which masked and minimized the actual brightness differences between the EBBs and the other sites. A more appropriate comparison would have been directly in front of each of the signs of interest so that the authors could be sure that they were comparing sign to sign without the contribution of the general ambient environment. Further, the authors do not state whether some of the (non-EBB) sites measured at night were those on surface streets, and whether there were fixed luminaires within the range of the luminance meter at such sites. The presence of fixed lighting would also have reduced the actual luminance differences between EBBs and other sign sites. Despite these limitations in measurement strategy, and despite the fact that the digital billboards were automatically dimmed at night, the authors recorded nighttime luminance levels at the driver's eye position that were, on average, 10 times greater for digital billboards than for baseline sites, approximately 3 times brighter than sites with conventional billboards, and approximately 2.5 times brighter than comparison sites. The authors' state: "this probably explains some of the driver performance findings in the presence of the digital billboards" (p. 68).

e. Different route types and control of variables

The authors conducted their on-road studies on "interstate, downtown, and residential road segments" (p.. 27) Given that all five (5) digital billboards (study sites) were on interstate highways, the decision to include some of the control sites (baseline, conventional billboards, comparison sites) on roads other than interstates confounded the data collection and made meaningful comparisons across sites impossible. When conducting field research, the goal must be to reduce, wherever possible, extraneous sources of variability. In this study, the decision to include study sites (EBBs) on interstates and some (we are not told which or how many) control sites on surface streets leads to additional uncontrolled sources of variability. Some of the significant differences between these two classes of roadways, any or all of which may have affected the data, are: traffic speeds and flow; lighting levels; sight distances; access control; at grade vs. grade separated intersections; traffic control devices; and divided vs. undivided traffic.

Even for the five EBBs that were the principal focus of this research, the authors seem to have made no attempt to identify, no less control, extraneous variables such as traffic speeds and volume, horizontal and vertical curvature, or other roadway and traffic characteristics that might have interacted with the variables of interest. Further, the distance between adjacent study sites was often very short. For example, using the Haversine formula, we calculated the distance between Site 37, an EBB, and Site 36, a baseline site, as less than 1.2km. Other studied sites might have been even closer to one another. Thus it is likely that the visibility ranges for adjacent sites overlapped, confounding eye gaze and vehicle performance comparisons.

f. Precision of eye gaze recording

The equipment and methodology used for recording eye gaze of the participant drivers was apparently chosen because it is more "naturalistic" (i.e. less intrusive for the participant). The downside of this choice is that this approach yields data that are less reliable and less accurate than would have been attained using more sophisticated equipment. As a result of the speed of vehicle movement, the distances of the vehicle from the EBBs and other sites of interest, the lack of careful calibration of the equipment prior to each run, and, especially, the vague definition of site boundaries, it is highly unlikely that the eye gaze data yields results with sufficient precision that enables the researchers to know whether any given glance was made to a site of interest, another vehicle, or another site that was not specified for data collection. Eye glance location of one degree of accuracy or better is probably necessary for a study where accuracy is critical; but the equipment and method used by these researchers most likely has accuracy of 10 degrees or worse.

g. Other methodological issues

The authors describe this project as a "naturalistic" driving study, modeled after a much larger and broader-based study conducted at the same institution – generally known as the "!00 car study" (Dingus, et al, 2006). Although they used an instrumented vehicle with on-board cameras, and although their test subjects drove the route without a researcher present in the vehicle, the present study differs significantly from the 100 car study in several key ways. First, the four on-board cameras used to record views of the road and of the drivers' glances were not hidden as they were in the 100 car study. Rather, they were prominently located on the driver's side A-pillar and on and adjacent to the rear view mirror. These camera locations are shown in Figures 8-10 of the report (pp. 32-33). Second, the duration of the present study was typically less than two (2) hours, whereas, in the 100 car study, participants kept their instrumented vehicles in their possession and used them daily for months. Third, participants in the present study had to follow a prescribed route, with instructions taped to the dashboard, whereas in the 100 car study, participants were free to drive when and where they chose in the course of performing their activities of daily living. In short, whereas the participants in the 100 car study may well have become acclimated to their test vehicles over time and ignored the fact that they were participating in a research study, the participants in the current study were fully aware that their performance and behavior was being monitored and recorded – thus their behavior could not reasonably be described as "naturalistic."

The authors report that, for each participant, they calibrated the eyeglance equipment in a hotel parking lot after the participant had driven the route. Given that eyeglance recording equipment can "drift" over time, that vibration could have changed the mounting position of one or more cameras, or that the driver could have adjusted the seat or otherwise changed position, this calibration should have been performed, for each participant, both before and after their drive. Without this comparison, the researchers had no way of knowing whether any recording parameters may have changed during the run.

3. Review and Application of Cited Literature

There is a long history of published literature examining the relationship of roadside billboards to crashes and to driver behavior. Relevant studies dating as far back as 1934 have been identified and reviewed by others. And research continues to be conducted and reported to the present day. The authors chose to discuss only a small, non-representative subset of studies. As will be seen below, it is clear that the studies reported, particularly of the early work in this field, were selected because they were supportive of the authors' position. When they cite studies that reported findings at odds with their position, these authors dismiss them as poorly done or irrelevant; conversely, studies that report findings consonant with those of these authors are praised with inappropriate descriptors such as "rigorous." The authors report at length about their own previous research study conducted with conventional billboards in North Carolina. That project used the same basic methodology as the current effort, with the same inherent flaws, and its findings are equally suspect.

Their reporting about two early epidemiological studies is illustrative of their approach to the literature.

The authors cite an article by Rykken (1951), a two-page interim progress report on a roadside study conducted in Minnesota. They quote from Rykken: "...no apparent relationship was found between accident occurrence and advertising sign type or location" (p. 12). What they do not say, however, is that Rykken called this result "a very preliminary study of approximately 170 mi. of the 500 mi. study segment (p. 42). Significantly, they fail to cite the final report of the subject study (Minnesota Department of Highways, 1951) which concluded, in part: "An increase in the number of advertising signs per mile will be accompanied by a corresponding increase in accident rate" (p. 31), and "intersections at which four or more (advertising) signs were located had an average accident rate of approximately three times that for intersections having no such signs." This final report has been extensively cited and reviewed by previous researchers. Wachtel and Netherton (1980), in particular, discussed it at length. It is puzzling, therefore, why these authors cited the interim progress report and ignored the final document.

The researchers followed the same approach in their review of a parallel study conducted in Michigan. They cite an interim study report by McMonagle (1951) that looked at only partial findings (p. 12), and ignored the study's final report (Michigan State Highway Department, 1952) which found that illuminated advertising signs showed "an appreciable association with accident locations" (p. 6).

In a confusing examination of a study by Rusch (1951) which analyzed crash reports on Federal and State highways in Iowa, the authors fail to report on Rusch's published results, and offer no evaluation of his study. Instead, they cite a brief review by Andreassen (1985) [ignoring other independent review of the Rusch work] which stated, in part: 'the greatest number of inattention accidents occurred on the sections where business and advertising predominated as the roadside property usage, but this does not prove anything about the effect of advertising signs on accident occurrence" (p. 13). Given that Rusch's findings, despite methodological weaknesses that often affected these early field studies, demonstrated that the number of accidents was more than double in the study section (where 90 percent of the businesses and roadside advertising signs were located) than in either of the two control sections, given that "inattention" accidents predominated over both "business" and "other" accident categories in this study section, and given that the results held when corrected for mileage per segment, the researchers' treatment of this study is puzzling.

Indeed, there were numerous examples of bias present in the authors' literature review. The four cited here are merely illustrative.

4. Statistical Methods And Analyses Used

a. Long duration eye glances

A major weakness of this study is the authors' failure to follow the very guidance that they recommend in their review of the work by Wierwille (1993), Horrey and Wickens (2006), and the "100 car study," (Dingus, et al., 2006), and their resultant misinterpretation of their own data. This failure is exacerbated by their decision not to present certain key data. Because this issue is central to their findings as well as to their apparent bias, it is discussed more fully in the paragraphs below.

As early as the Abstract the authors state: "Various researchers have proposed that glance lengths of 1.6 seconds, 2.0 seconds, and longer may pose a safety hazard. An examination of longer individual glances showed no differences in distribution of longer glances between the four event types" (p. 6). They restate this conclusion on page 9: "An analysis of glances lasting longer than 1.6 seconds showed no obvious differences in the distribution of these longer glances across event types." As will be seen below, this is simply not true.

In their introductory description of eyeglance results (p. 52) the authors list the seven (7) questions that they sought to answer. The seventh was: "Are longer glances (longer than 1.6 s) associated more with any of the event types?" This is followed by a summary and analysis of the findings relative to these seven questions. In all cases except one, the researchers performed an analysis of variance (ANOVA) to analyze the data, and reported their tests of statistical significance in both graphical and narrative form. The one exception was with regard to Question 7, longer glance durations. Despite restating Wierwille's recommendation that 1.6s be used as a criterion representing a long glance away from the roadway, and despite explaining that their approach in analyzing this data follows that provided by Horrey and Wickens, "who suggest analyzing the tails of the distributions whenever eyeglance analysis is performed" (p. 59), these authors failed to do any such analysis. Instead, they apparently performed only a visual inspection of the data contained in their Figure 23 (p. 59) which depicts the distribution of glance durations for the four different event types. As a result of this visual inspection, they state: "As shown in Figure 23, the distributions of glance duration were similar across all event types, and there was no obvious pattern of longer glances being associated with any of the event types" (p. 59).

Figure 3, below, reproduces the authors' Figure 23 together with its original caption.

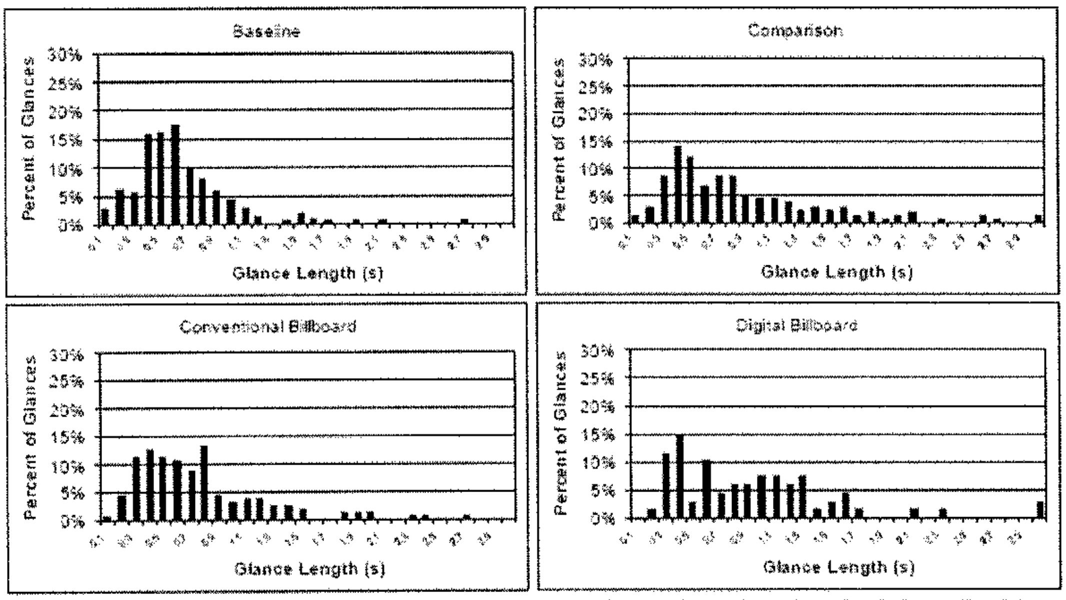

Figure 23. Tails analysis for the distribution of glance duration, (method described in Horrey and Wickens, 2007).

Figure 3. A reproduction, in original size, of the authors' Figure 23, together with its original caption.

Because the authors do not provide data on the actual glance durations that were used to prepare these four bar charts, the reviewer cannot perform an independent review of their conclusions. However, enlarging the four charts enables at least a rudimentary independent review of their findings, with the following results. Using only the tails analysis as suggested by the authors (following Horrey and Wickens), and using both 1.6 seconds (the Wierwille criterion) and 2.0 seconds (based on findings from the 100 car study), we find the following:

> For glances of 1.6 seconds or longer, approximately 5.5% of baseline sites and 7.5% of conventional billboard sites captured glances of such long duration, whereas 13% of digital billboards and 16% of comparison sites met or exceeded this criterion.

> For glances of 2.0 seconds or longer, approximately 2% of baseline sites and 4.5% of conventional billboard sites captured glances of such long duration, whereas 7% of digital billboards and 8% of comparison sites met or exceeded this criterion. Indeed, both EBBs and comparison sites captured glances longer than 3 seconds, behaviors not experienced with either baseline or conventional sites.

Given that this visual inspection of the researchers' data suggests that long glances seem to occur two-three times more often with EBBs and comparison sites than they do with baseline or conventional sites, such results suggest important differences for the longest glances, the ones that highway safety experts are most concerned with. One must ask why the authors chose not to perform a statistical analysis of this data, particularly when they did so for every other set of eyeglance data, why their visual inspection of these data did not identify the dramatic differences reported here

b. Findings of the nighttime study

In the authors' abbreviated nighttime data collection effort, a shorter route with fewer signs and fewer participants, they examined four different eyeglance measures – eyes-on-road percent, overall glance frequency, mean glance duration in the direction of an event, and mean number of glances in the direction of an event. In *all four cases* they found that EBB and comparison sites showed differences from conventional billboard and baseline sites – always in the direction to suggest that the EBBs and comparison sites were more attention-getting and more distracting (pp. 64-66). Surprisingly, especially given their emphasis on longer glance durations based on the "100 car study" and the work by Wierwille, they apparently made no attempt to capture data for long glance durations, no less report it.

c. Obfuscation of correlation with causation

Throughout the report, the authors confuse the terms *correlation* and *causation*. Although it is clear that they well understand the important differences between these two types of statistical analysis, they often slip into the erroneous mode of citing a study whose sole purpose was to measure correlation, and criticize that study because it failed to prove causation. These fallacious comments are in line with a long tradition in the outdoor advertising industry of suggesting that there can be no relationship between billboards and traffic safety because billboards have never been shown to *cause* accidents.

5. Misleading and Inconsistent Reporting, and Evidence of Bias

Beginning with the first sentence of the Abstract, and continuing throughout the report, the authors construct their language in such a way as to intentionally mislead the reader.

This was *not* a "naturalistic" study as stated in the Abstract (see our discussion of Methods) and it should not be reported as such.

The stated purpose of this study was to "assess the effects, if any, of digital billboards on driver behavior and performance" (p. 8), *not,* as implied in the Abstract, to ascertain whether driving performance in the presence of digital billboards was on a par with driving performance in the presence of on-premise signs that were also bright, electronic, and changing. The researchers clearly found that EBBs *did* have an adverse impact on driving performance, and the fact that this adverse impact was similar to the adverse impact from other distracting signs does not diminish this finding nor make it acceptable. The authors admit that "there are measurable changes in driver performance in the presence of digital billboards" (p. 6), and, as demonstrated in the body of the report, these changes are adverse, but such adverse impacts cannot be deemed acceptable merely because they "are on a par" with the adverse effects of other digital, flashing, and changing signage located on the premises of roadside businesses.

Baseline sites should have been, as stated in the abstract, "sites with no signs." But, as described elsewhere in the report, an unidentified number of them *did* contain signs, thus diminishing their potential to serve as intended control sites.

In direct conflict with a statement in the Abstract, longer individual glance patterns (greater than 1.6 and 2.0 seconds) *did* show differences (actually, rather dramatic differences) between the event types. In fact, per the authors' own statements elsewhere in the report, these differences at the tails of the distributions for glance duration may be critically important in assessing the true impact of digital billboards on driver performance and behavior.

Similar misstatements are made throughout the Executive Summary, and will not be repeated here. However, the expressed "finding" that: "An analysis of glances lasting longer than 1.6 seconds indicated that these longer glances were distributed evenly across the digital billboards, conventional billboards, comparison events, and baseline events during the daytime" (p. 7) is plainly false, as is discussed at length in our section dealing with Statistics. Significantly, the data discussed in this "finding" was not analyzed by the researchers in accordance with their own data analysis recommendations, nor was such data even collected for the abbreviated nighttime study, when we would have expected such findings to be even more dramatic than they were in the daytime study.

The authors identified five (5) EBBs for study. These are identified by latitude, longitude, route number, and side of road in Table 2, and shown graphically on a map in Figure 2. With this information, that reader can view images of these

EBBs from either the Tantala report or from the website of Clear Channel Outdoor, at http://www.clearchanneloutdoor.com/products/digital/don/cleveland/index.htm . Examination of Figures 1 and 2 in our report, must lead the reader to question the accuracy of the authors' statement on p. 19: "The Cleveland digital billboards...were located off to the side of the roadway in straight-away sections of interstate with no interference from hills, curves, or intersections."

The authors provide voluminous data for irrelevant issues (e.g. 124,740 video frames analyzed, 96,228 data points collected, 8,678 eye glances identified, etc.) but offer no information useful to readers who might want to know what was actually studied (for example, there are no images of any of the billboards or other sites studied, there is no indication of the precision with which eye gaze was captured, etc.). It appears as if the researchers intended to overwhelm the reader with useless information in an attempt to avoid questions about the real issues.

There are numerous statements throughout the report that, on the one hand, are irrelevant to the study, and, on the other, demonstrate a clear pro-billboard attitude. Some examples:

> "The lead author of this report recently participated on an expert panel charged with providing recommendations for a minimal data set to be included on police accident reports; billboard were never raised as a possible distraction..." (p. 11).
>
> "After a long gap in research, there were a few additional studies in the 1960's through the 1980's, none of which demonstrated that billboards were unsafe." (p. 11)
>
> "The national crash databases do not mention billboards in their list of driver distractions." (p. 14)

The authors have coined the term *safety neutral* (p. 10). They state: "Although there are measurable changes in driver performance in the presence of digital billboards, in many cases these differences are on a par with those associated with everyday driving, such as the on-premise signs located at businesses" (p. 6). In other words, the authors seem to be saying, because other roadside distractions such as their "comparison sites" (which, the authors note elsewhere, may contain multiple signs, changeable message signs, and digital signs) may also attract drivers' attention and are associated with speed and lateral placement difficulties as well as long glances away from the forward roadway, EBBs should be considered *safety neutral* because their adverse effects are little different than these other sources of distraction..

The authors continually obfuscate the difference between correlation and causation. In reporting on a study by Garvey, et al. (1995), they state: "the common problem with these studies is attributing accident causation..." (p. 13). Given that the studies quoted generally made no attempt to establish causation, this is not a problem with these studies, but rather an accusation by these authors that such studies claimed to be something other than they actually addressed. Similarly, in their review of work by Andreassen (1985), they state: "almost all studies have relied on correlations and/or subjectively assigned 'inattention' factors, which can only produce very tenuous evidence for a causal link between advertising and accident frequency" (p. 13).

As discussed earlier in this report, there was a serious confounding in the visual and attention-getting characteristics of the EBBs which were the subject of this study, and the "comparison sites" which were used as one of the three types of control sites. This confounding was due directly to the researchers' choice of such comparison sites which, in their own language, often included digital signs, changeable message signs, and flashing signs. That the researchers' often found quite similar driver performance and behavior in these two types of sites, and that these performance and behavior variables typically differed from the two other types of sites studied (conventional billboards and baseline sites) should have sent a clear signal that sites containing digital imagery with changing messages were more demanding and more distracting than sites devoid of such sign characteristics. Yet, the authors took this obvious conclusion and twisted it in favor of their biases by simply assuming that the "comparison sites" formed part of the normal driving environment and, therefore, that the presence of EBBs was no worse than this normal environment and therefore was no problem. This conclusion demonstrates obvious bias, and flies in the face of efforts to promote highway safety by reducing, not increasing, the number of irrelevant, distracting, roadside stimuli. (By way of analogy, if we know from prior research that the use of mobile telephones in vehicles contributes to driver inattention, and we find that the use of some newer technology leads to a similar degree of inattentiveness, we would be hard pressed to find this newer technology acceptable merely because it is no worse than the distraction from telephone use).

F. IMPLICATIONS FOR POLICY.

To the best of our knowledge, this is the first independent peer review of these two reports to have been prepared. There may be others underway or planned, and the authors of the two reports may well wish to respond to any such peer review. Because of the public relations campaign with which the OAAA released and publicized these two studies, they have received wide press coverage in print, online, and in the broadcast media. Without exception, this coverage has presented uncritical acceptance of these two reports as presented, with no scrutiny of their scientific or technical soundness. As a result, numerous States and local government agencies have begun to modify their codes and ordinances that address the use of digital billboards along the roadside. Having completed this peer review, it is our opinion that acceptance of these reports as valid is inappropriate and unsupported by scientific data, and that ordinance or code changes based on their findings is ill advised. Even the Federal Highway Administration (FHWA) has issued a recent policy memorandum in which DBBs are given tacit acceptance under certain conditions, possibly based in part on the release of these two studies. Because FHWA remains concerned about the safety implications of EBBs on highways, and because of its stated intention to conduct or sponsor its own research into this issue, it seems to this writer logical that any such policy change await further developments from research.

This reviewer, after careful review of different research methodologies that have been undertaken or suggested to answer these questions, recognizes that the design and conduct of such research is challenging, and that no research yet published can fully answer the question of whether EBBs create a sufficient distraction of drivers' attention that they should be banned or strictly regulated under certain roadway, environmental, and traffic conditions. Such research is needed to guide future policy and regulation.

At the forthcoming 87th Annual Meeting of the Transportation Research Board, to be held in Washington, DC starting on January 13, 2008, this reviewer will conduct a day-long workshop titled: "Digital Billboards on the Highway: A Bright Future?" This workshop will provide a forum for a discussion of the complex issues involved in this field of inquiry. In addition, this reviewer will hold the first meeting of a new TRB Subcommittee on this topic during the TRB meetings, and interested parties are invited to attend.

Finally, it is suggested, given the interest in this subject expressed by many State and local government agencies, that interested States seek funding to support the conduct of an objective, independent research project at the earliest possible time.

G. REFERENCES

American Association of State Highway and Transportation Officials (AASHTO) (2007). *Proposed Policy Resolution – Outdoor Advertising.* Highways Subcommittee on Traffic Engineering.

Andreasen, D.C. (1985). Technical Note No. 1: Traffic Accidents and advertising signs. *Australian Road Research Board,* 15(2), 103-105.

Dingus, T.A.., Klauer, S.G., Neale, V.L., Petersen, Al, Lee, S.E., Sudweeks, J., Perez, M.A., Hankey, J., Rasey, D., Gupta, S., Bucher, C., Doerzaph, Z.R., Jermeland, J., & Knipling, R.R. (2006). *The 100-Car Naturalistic Driving Study: Phase II – Results of the 100-Car Field Experiment.* Report No. DOT HS 810 593. Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Farbry, J., Wochinger, K., Shafer, T., Owens, N., & Nedzesky, A. (2001). Research review of potential safety effects of electronic billboards on driver attention and distraction – Final Report. Washington, DC: U.S. Department of Transportation, Federal Highway Administration.

Horrey, W.J. & Wickens, C.D. (2007). In-vehicle glance duration: Distributions, tails and a model of crash risk. *Proceedings of the 2007 Transportation Research Board Meeting.* Washington, DC: TRB.

Massachusetts Outdoor Advertising Board (1976). *Decision in the matter of John Donnelly & Sons, Permittee, Telespot of New England, Inc., Intervener, and Department of Public Works, Intervener, with Respect to Permit Numbered 19260, As Amended.* Boston, Massachusetts.

McMonagle, J.C. (1951). Accident analysis – Telegraph road 1947-1948. *Highway Research Board Bulletin,* 30, 29-41.

Michigan State Highway Department (1952). *Accident Experience in Relation to Road and Roadside Features.*

Minnesota Department of Highways (1951). *Minnesota Rural Trunk Highway Accident, Access Point and Advertising Sign Study.*

Outdoor Advertising Association of America (2007). Ground-breaking Studies Determine Accidents Not More Likely To Occur Because of Digital Billboards. http://www.oaaa.org/ (July 11, 2007).

Rykken, K.B. (1951). Minnesota roadside survey: progress report on accident, access point and advertising sign study in Minnesota. *Highway Research Board Bulletin,* 30, 42-43.

Rusch, W.A. (1951). Highway accident rates as related to roadside business and advertising. *Highway Research Board Bulletin,* 30m 46-50.

Shepherd, G.M. (2007). *Memorandum – Guidance on Off-Premise Changeable Message Signs.* Washington, DC: U.S. Department of Transportation, Federal Highway Administration.

Slobodzian, J.A. (2007). "Video billboards coming this way – Police praise them. Conservationists and driver-safety groups oppose them." *Phily.com* (Obtained from the internet at http://www.philly.com/philly/business/9278261.html)

Wachtel, J., & Netherton, R. (1980). *Safety and Environmental Considerations in the Use of Commercial Electronic Variable-Message Signage (Final Report).* Report No. FHWA/RD-80-051. Washington, DC: U.S. Department of Transportation, Federal Highway Administration.

Wierwille, W.W. (1993). Visual and manual demands of in-car controls and displays. In B. Peacock and W. Karwowski (Eds.)., *Automotive Ergonomics* (pp. 307-316). Amsterdam: North Holland Press.