

Self-Regulation Minimizes Crash Risk from Attentional Effects of Cognitive Load during Auditory-Vocal Tasks

Richard Young
Wayne State Univ.

ABSTRACT

This study reanalyzes the data from a recent experimental report from the University of Utah investigating the effect on driving performance of auditory-vocal secondary tasks (such as cell phone and passenger conversations, speech-to-text, and a complex artificial cognitive task). The current objective is to estimate the relative risk of crashes associated with such auditory-vocal tasks. Contrary to the Utah study's assumption of an increase in crash risk from the attentional effects of cognitive load, a deeper analysis of the Utah data shows that driver self-regulation provides an effective countermeasure that offsets possible increases in crash risk. For example, drivers self-regulated their following distances to compensate for the slight increases in brake response time while performing auditory-vocal tasks. This new finding is supported by naturalistic driving data showing that cell phone conversation does not increase crash risk above that of normal baseline driving. The Utah data are next compared to those from a larger study that included visual-manual as well as auditory-vocal tasks. The Utah auditory-vocal tasks had negligible effects on response time compared to visual-manual tasks with socially acceptable crash risk, such as manual radio tuning. In conclusion, auditory-vocal tasks such as those in the Utah study are not expected to increase crash risk or impair real-world operation of an automobile, compared to normal baseline driving without performing such tasks.

CITATION: Young, R., "Self-Regulation Minimizes Crash Risk from Attentional Effects of Cognitive Load during Auditory-Vocal Tasks," *SAE Int. J. Trans. Safety* 2(1):2014, doi:10.4271/2014-01-0448.

OUTLINE

Part 1 reviews a recent experimental study of auditory-vocal tasks conducted by researchers at the University of Utah [1] (hereafter called the "Utah" study). Its "Cognitive Distraction Index" (hereafter called the "Index") for auditory-vocal tasks is reviewed. In Part 1, the original data for key variables making up this Index are re-analyzed. The Index scores are compared to the real-world relative crash risk for cellular conversations in naturalistic driving. Part 1 examines whether the Utah data support its assumption that auditory-vocal tasks increase crash risk, or, support the alternate view that driver self-regulation maintains crash risk at normal driving levels during auditory-vocal tasks. Part 2 compares key data for the Utah auditory-vocal tasks to the Crash Avoidance Metrics Partnership Driver Workload Metrics (CAMP-DWM) data [2] for auditory-vocal and visual-manual tasks.

PART 1. THE UTAH STUDY

1.1 Introduction

1.1.1 Utah Cognitive Index

The Utah study used a large number of variables and test conditions to assess the attentional effects of cognitive load from auditory-vocal tasks on driving. It tested 7 auditory-vocal

task conditions and 1 "single" (no secondary task) condition in 3 experimental venues: standalone with no driving, in a simulator while following a lead vehicle, and in an instrumented vehicle in on-road city traffic at about 25 mph. It gathered 31 dependent variables, condensed into 13 final variables by averaging the scores for the same variable collected in different venues. From this condensed set of 13 variables, the Utah investigators created its Index. Subjective workload scales constituted 6 of the final 13 variables and so were heavily weighted in the Index. The 13 variables were "standardized" to a common scale, and summed to create a single combined score, which was transformed into a 5-point scale by anchoring the low end at "Category 1" with the composite score for the "Single" baseline condition with no secondary task. The high end was anchored at "Category 5" with the composite score for a complex artificial task called "OSPAN" in which subjects attempted to recall single syllable words in serial order while solving mathematical problems [1, p. 13]. Figure 1 re-plots the final scores of the 8 task conditions as given in [1, Figure 28].

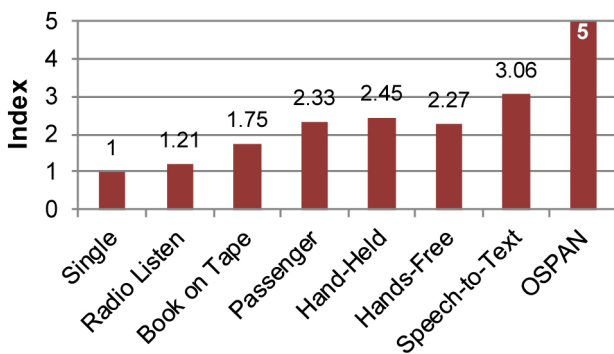


Figure 1. Utah results for 8 auditory-vocal task conditions. The scores on the Utah Index range from 1 to 5 (re-plotted from [1, Figure 28]).

1.1.2 Crash Risk Assumption in the Utah Study

The Utah study [1, p. 30] assumes that its Index relates to crash risk, "It is reasonable to assume that there would be a monotonic relationship between cognitive distraction and crash risk." (Although not explicitly stated, it is obvious that the Utah report assumes a monotonic *increase*, not a monotonic *decrease*, in crash risk with increased cognitive distraction.)

The Utah study placed its three conversation tasks - talking to a passenger, talking on a hand-held phone, and talking on a hands-free phone (Figure 1, middle bars) - into a "moderate" Category 2 cognitive demand level. The study concludes that conversation tasks, "...are significant impairments to driving that stem from the diversion of attention from the task of operating a motor vehicle." [1, p. 30]

Another task called "speech-to-text" involved a simulated voice interaction with a machine rather than a human. The Utah study concludes that this task (with a score of 3.06; second-to-last bar in Figure 1) was even more cognitively demanding than conversation, and so placed it in a higher "Category 3" Index level. The report states [1, p. 29]:

Given the current trends toward more voice commands in the vehicle, this Category 3 level of cognitive distraction is troubling. The assumption that if the eyes were on the road and the hands were on the steering wheel then voice-based interactions would be safe appears to be unwarranted.

The study concludes [1, p. 30] that adoption of voice-based speech-to-text and text-to-speech systems in the vehicle, "... may have unintended consequences that adversely affect traffic safety."

The report identified what it considers some potential driving "costs" incurred with doing the auditory-vocal tasks it tested, and attempted to reflect these "costs" in its Index (Figure 1). It refers 13 times to these costs as "impairments" to the driver, which in some cases "may rise to the level associated with drunk driving" [1, p. 29]. The use of a 5-category scale deliberately modeled on the Saffir-Simpson hurricane scale implies that safety concerns approach catastrophic levels at the upper end of the scale. The average reader is left with the

impression that auditory-vocal tasks performed while driving increase crash risk by enormous amounts, even to the level associated with drunk driving [1, p. 29].

However, there were no reported crashes on the road or in the simulator in the Utah study, and no mention of the number or percentage of experimenter interventions to alert the drivers to the presence of a roadway hazard on the road, belying the implication that a Category 5 distraction has enormous absolute crash risk. There was no evidence presented comparing the Index to known major impairments, such as drowsiness, drugs, or alcohol. No visual-manual tasks were tested for comparison to the auditory-vocal tasks. Indeed, prior experimental studies [2, 3] found that speech interfaces were less distracting than visual-manual interfaces. Without such a context, the upper bound of the Utah scale is unknowable. A higher upper bound would dramatically affect the relative significance of what is being expressed by the categories on the Utah scale. Hence, the validity of the Utah study's assumptions, data, results, and conclusions about its Index bear scrutiny.

1.1.3 Driver Self-Regulation

An important fact not recognized in the Utah study [1] is that drivers routinely self-regulate their behavior during real-world driving. Drivers can, and do, "actively adjust their driving behavior in response to change or competing task demands to maintain an adequate level of safe driving" [3, 38]. Such self-regulation can occur from drivers choosing not to do a secondary task while driving, or adjusting their driving behavior while performing the task to keep their safety margins equal to or better than when not doing the task. Self-regulation is typically not observed in experimental studies, because drivers are not given a choice about performing a task (i.e., they must perform the tasks to meet the requirements of the experiment).

Many real-world non-experimental studies have observed such self-regulatory behavior in everyday driving, and proposed that drivers self-regulate their driving to compensate for the attentional effects of cognitive load and thereby reduce crash risk. For example, the *compensation hypothesis* was proposed by Young and Schreiner [4] to explain their finding that using the OnStar embedded device to make hands-free personal calls does not elevate the rate of crashes severe enough to deploy an airbag [4]. The compensation hypothesis states that drivers talking on a cell phone during natural driving tend to change their driving behavior such that there is no net safety decrement from the slightly increased response time (RT) to events during such conversations that is predicted from experimental studies [5, 6]. Indeed, there is previously unrecognized evidence of such self-regulation in the Utah study data, as shown in the current study.

1.2 Method

The means and standard errors for the “response time” and “following distance” variables were requested and received from the lead author of the Utah report [7]. These original data, rather than the “standard” scores in Appendix A of the Utah report [1], were used for the current in-depth analysis, because the “standard” scores were actually not standard scores as claimed in the Utah report. Across the 8 task conditions, each variable in its Appendix A has a mean of zero, but the standard deviations are not 1, as required for a standard score. The current study therefore used only the original data to investigate key variables in the Index.

1.3 Results

1.3.1 Brake RT to Brake Light Onset

Figure 2 plots the subject driver mean brake response time (RT) to the pace vehicle brake light onset in the Utah simulator experiment, reproducing Figure 23 in the original report [1]. Note that the brake RT in Figure 2 varies in the same manner as the Index scores in Figure 1 ($r = 0.95$, $p = 0.0004$). That is, the higher the brake RT, the higher the Index. This result is not surprising, since brake RT was one of the 13 variables used in the Utah report to create its Index. The largest RT was for the OSPAN task (1092 ms), which is a statistically significant increase of 171 ms, compared to the baseline “Single” condition (921 ms) ($t = 4.36$, $p = 0.00005$, independent two-sample t-test, equal sample sizes, equal variance). The increases for the three conversation tasks in the center of Figure 2 were 26 to 56 ms compared to the Single baseline (not statistically significant). These RT increases are near the low end of the 50-200 ms range (about the time of an eye blink) found in prior experimental venues (brain imaging, laboratory, simulator, test track, and open road), for the effect of conversation tasks on visual event RTs [2, 5, 6].

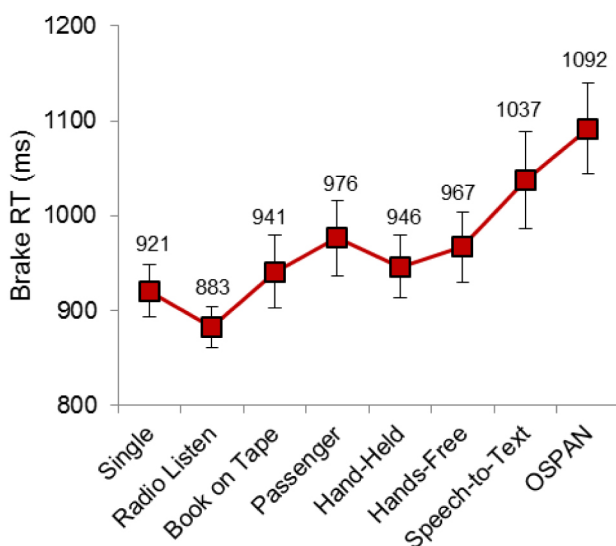


Figure 2. The mean brake RT to a lead vehicle brake light onset for the 8 task conditions in the Utah simulator. Error bars are standard errors.

1.3.2 Following Distance to Pace Vehicle

Figure 3 plots the following distances in the Utah simulator experiment for the 8 task conditions. Note that the following distance also varies in a manner similar to the Index score in Figure 1 ($r = 0.92$, $p = 0.0014$). That is, the higher the Index, the greater the following distance. This result is again not surprising, since following distance was another of the 13 key variables on which the Index was based.

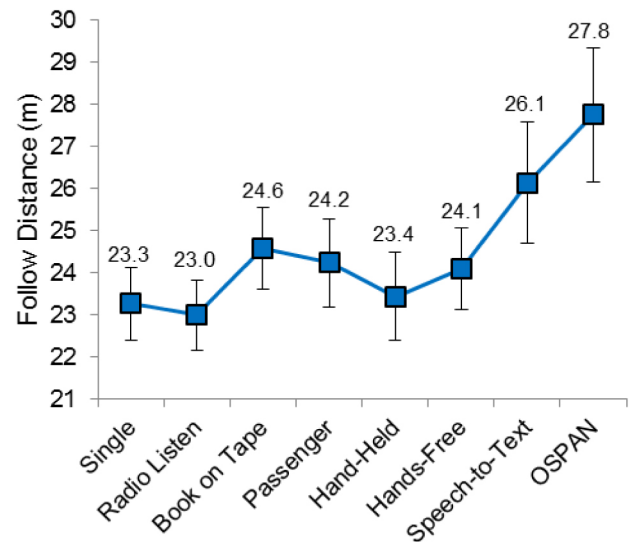


Figure 3. The mean following distance to the lead vehicle for the 8 task conditions in the Utah simulator. Error bars are standard errors.

1.3.3 Increased Following Distance Compensates for Increased Brake RT

Previous studies [19, 20, 21, 22, 95], including those of the Utah investigators [11, 12, 13, 14, 15, 16, 17, 18], have interpreted increased following distance as driver self-regulation, providing an increased safety margin while performing auditory-vocal tasks. Should the lead vehicle suddenly apply its brakes and decelerate rapidly, increased following distance gives subjects more time to brake, thereby avoiding increased rear-end crash risk. The brake RT and following distance are well correlated ($r = 0.96$, $n = 8$, $p = 0.0002$), as is self-evident by comparing Figures 2 and 3.

Do these increases in following distance compensate for the increases in brake RT? The Utah study claims they do not: “Brake RT increased as a function of condition over and above any compensatory effects associated with following distance” [1, p. 18]. The basis for this conclusion was a “subsidiary linear mixed model analysis that held following distance constant,” but details were not provided of this model, its results, or why holding following distance constant would indicate it was not a compensatory behavior. Figure 3 shows that following distance was actually not constant in the Utah data, raising further questions of why the Utah model would hold it constant. A closer examination of following distance and its potential compensatory effects is therefore warranted.

One way to test if the observed increases in following distance (meters in [Figure 3](#)) were sufficient to compensate for the observed increases in brake RT (milliseconds in [Figure 2](#)) is to place the following distance on the same time scale as brake RT. This conversion is easy if the speed of the vehicle were known, because distance can then be transformed into time. However, the Utah investigators did not report the mean speed of the subject vehicles during the approximately 10-min task trials in the simulator. Therefore, transformation of the Utah study following distance to a time scale cannot be done with available information. (Reasonable speed assumptions for the Utah study give rise to similar results as the current study [8].)

In lieu of speed data, the compensation question can still be directly assessed by calculating the percentage *changes* in brake RT and following distance relative to the baseline “Single” condition. Speed affects only the *absolute* time headway; the *relative* time headway equals the relative distance headway, regardless of the average speed. For example, consider the following equations showing following distance D , speed S , and the headway time T for two tasks:

$$\begin{aligned} D_1 &= S_1 T_1 \\ D_2 &= S_2 T_2 \end{aligned} \quad (1)$$

Taking the ratio of [Equations 1](#) gives:

$$D_1 / D_2 = (S_1 T_1) / (S_2 T_2) \quad (2)$$

The pace car had the same average speed for every task condition [9]. Otherwise, the driver performance metrics would have been confounded with speed in the experimental design, invalidating the Utah results. In addition, subjects would have tried to match their speed to the pace car as best they could in order to keep a fixed gap to the pace car as instructed. Hence, $S_1 = S_2$ and speed cancels out of [Equation 2](#), leaving the ratio of the following distances equal to the ratio of the headway times:

$$D_1 / D_2 = T_1 / T_2 \quad (3)$$

[Equation 3](#) indicates that the ratio of the following distances D is equal to the ratio of the headway times T for the Utah simulator experiment. It follows that if the following distance increased proportionately to the brake RT for a given task (relative to the Single condition with no secondary task), then the subjects must have compensated for their increased brake RT by increasing their headway time (which was proportional to their following distance).

To test this simple deduction, the data in [Figures 2](#) and [3](#) were each converted into percentages relative to their respective Single baseline. [Figure 4](#) plots the percentage change in following distance (blue bars) vs. that for brake RT (red bars) on the same y-axis percentage scale for comparison.

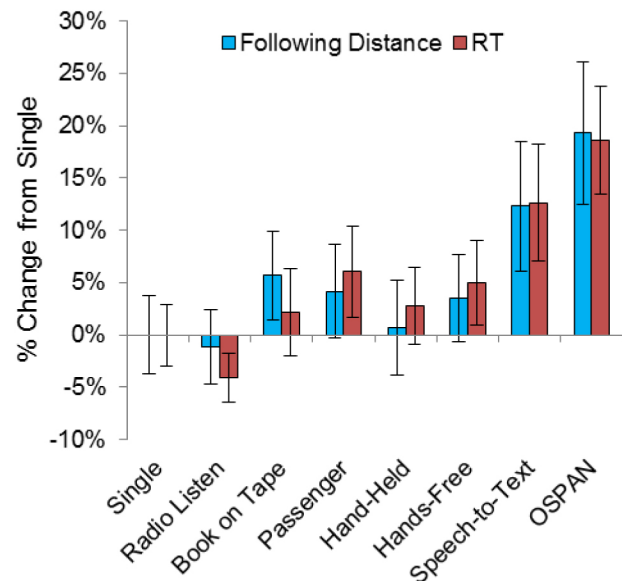


Figure 4. The percentage change relative to the Single baseline for following distance (blue) and brake RT (red) for the 8 task conditions. Error bars are standard errors.

For all 7 auditory-vocal tasks tested, [Figure 4](#) shows that the following distance changed in a proportional manner (relative to the “Single” baseline), as did the brake RT. Not one task had a percentage increase in RT that was significantly larger than the percentage increase in following distance. Hence, under the conditions of the Utah simulator study, subjects compensated for the increases in brake RTs with proportionate increases in their following distances for all tasks in the study.

For example, [Figure 2](#) shows that the RT increased from 921 ms for the “Single” control condition, to 1092 ms for the most difficult OPSAN task condition, a 171 ms or 18.6% increase (rightmost red bar in [Figure 4](#)). [Figure 3](#) shows that the following distance increased from 23.3 m for the “Single” control condition, to 27.8 m for the most difficult OPSAN task condition, a 4.5 m or 19.3% increase (rightmost blue bar in [Figure 4](#)). Because these percentage increases were about the same, the increased following distance (and corresponding time headway) fully compensated for the increased RT.

Note that the subjects increased their following distances despite being instructed and trained, “to maintain a two-second following distance [*sic*] behind the pace car” [1, p. 21]. Evidently, drivers' self-regulatory behavior was so strong that it even overcame the strong influence of “obedience to authority” (the instruction and training to maintain a constant 2-s following time for every task condition).

It is unknown if the increase in RTs from secondary auditory-vocal tasks in simulated driving would be large enough in real driving to affect driving safety, even in the absence of compensation from self-regulation. What Figure 4 demonstrates, however, is that the increased brake RTs in the Utah simulator study were fully offset by the increased following distances during those tasks, as evidenced by the Utah study's own data. Hence, the Utah study's data does not support its conclusion that the cognitive loads of auditory-vocal tasks cause "significant impairments to driving." In fact, the Utah data in Figure 4 shows no impairment to driving for any of the dual task conditions, relative to the Single task condition. Drivers maintained an adequate safety margin in all task conditions, due to self-regulation.

1.3.4 Estimating Crash Risk from Key Utah Index Variables

1.3.4.1 Brake Response Time and Crash Risk

The Utah study assumes [1, p. 30] that an increased Index score (of which brake RT is a key part) increases crash risk. That argument assumes the absence of compensatory adjustments from other factors. However, the specific relationship between crash risk and RT is not simple, and can be complicated by environmental conditions, road geometry, and many other factors [10]. For example, the brake RT of a following vehicle to a lead vehicle brake light onset is estimated from crash investigation data to be about 639 ms shorter when a driver is behind a lead vehicle in a curve vs. a straight road, despite the fact that crash risk is higher on curves vs. straight roads [10, Table 3]. That is, shorter brake RTs are associated with higher crash risk in curves, exactly the opposite of the Utah study monotonic increase assumption. This result is likely due to increased vigilance by drivers while negotiating curves, thus reducing their RTs, yet the crash risk is higher because of the more complicated road geometry.

1.3.4.2 Following Distance and Crash Risk

It is noteworthy that the Utah subjects increased their following distance despite the fact that they were in an experimental setting and had been trained and instructed beforehand to maintain a constant 2-s headway. The analysis in Figure 4 shows that the following distances of the drivers in the Utah study were sufficient to compensate for the increased RT for all tasks in the Utah study, even the most difficult "Category 5" OSPAN task. Increased time and/or distance headways during cell phone conversation or other auditory-vocal tasks is also evident in prior simulator studies by members of the Utah group [13, 14, 15, 16, 17, 18] and others [19, 20, 21, 22, 95], with only one known exception [23].

In naturalistic driving studies (NDS), drivers have no authority figure instructing them to maintain a constant headway. Indeed, drivers in NDS increase their time headway when conversing on three phone types: hand-held, portable hands-free, and integrated hands-free cell phone [24, Figure 18]. Fitch et al. [24, Table 94] reported that the mean time headway difference

score for conversation vs. no conversation for each of the three phone types individually had p -values approaching significance at 0.095, and claimed there was no statistically significant effect of conversation on time headway. However, the current study pooled the p -values across the three cell phone types using the meta-analysis program in [25], and, contrary to Fitch et al.'s conclusion [17], found that there was a statistically significant effect of conversation at increasing time headways ($p = 0.029$). Thus, the real-world NDS data are consistent with the result of the current re-analysis of the Utah simulator data - drivers increase their following distances during cell phone conversation.

Do such increased headways reduce or increase crash risk? Some might argue that increased headways are caused by drivers not being able to perform the primary driving task properly under conditions of cognitive load, and therefore indicate "impairment" of the driver, which presumably increases crash risk. However, increased following distance is known to *reduce* real-world crash risk and traffic violations. For example, Evans and Wasieleski [26] measured headways of individual vehicles in real-world observational studies and then used each license plate number to look up the record of each vehicle's crash and ticket involvement, and found that increased headways were associated with fewer crashes and traffic violations. Hence, increased following distance (i.e., the reverse of tailgating) is a safety benefit in real driving. It is therefore invalid and misleading for the Utah report to include following distance in an Index that is assumed to correlate with increased crash risk, when in fact increases in following distance reduce crash risk.

1.3.4.3 Reduced Head Turns at Intersections

Locations with stop signs, stoplights, and/or crosswalks at intersections along the 25 mph speed limit city route of the Utah open road experiment were identified *a priori* by subjective judgment of the Utah investigators as posing a "potential" hazard. The Utah study calculated the probability of a left-right head turn by the subject driver at an intersection (presumably to fully scan for crossing traffic or pedestrians at those locations). (if the subject driver turned to look in just one direction but not the other, the Utah study marked it against the subject.) This head-turn probability was estimated from the video recordings in post-analysis and entered as a variable into the Index. If the probability (that the driver turned his/her head in both directions to scan the intersection) decreased during a task, the Utah study took that as indicating an effect of cognitive distraction. During the Passenger, Hand-Held, and Hands-Free conversation tasks, the percentage of scanning both directions at the intersection was about 66%. For the Single (baseline) condition it was about 74%, and for the most difficult OPSAN task it was about 61% [1, Figure 26]. This slight reduction in left-right scanning is partially consistent with the simulator results of Harbluk et al. [27, p. 272], who found that when approaching and driving through intersections, drivers reduced their scanning of intersection areas in the right

Likewise, Gruzdaitis et al. [28] found a reduction in scans to the right at intersections in an urban on-road instrumented vehicle study with 15 younger subjects in Helsinki, Finland.

However, Fitch et al. [24, p. xxvii] in an NDS study of cell phone effects of driving performance state: "The likelihood of properly scanning an intersection was not found to differ when using a cell phone." They found that drivers properly scanned the intersection during 38 percent of the baseline periods, 36 percent of the grouped visual-manual cell phone subtasks, and 35 percent of the grouped talking/listening cell phone subtasks. These small differences in scanning behavior are inconsistent with the Utah [1], Harbluk et al. [27], and Gruzdaitis et al. [28] results. In addition, the percentages of head turns (35-38%) in the Fitch et al. NDS study [24] are about half the percentages of head turns in the Utah on-road experimental study (61%-74%), for all task conditions. That is, the slightly reduced Utah head-turn probability at intersections when performing auditory-vocal tasks was still far higher than that observed during naturalistic baseline driving [24]. It is therefore difficult to argue that the observed percentage of head turns was indicative of increased crash risk compared to naturalistic driving. It is more plausible that the Utah study drivers were constantly aware of being observed by the experimenters in the car with them, and so were motivated to perform in the manner that they thought the experimenters expected them to perform, giving rise to far more head turns than would ordinarily occur in non-experimental real-world driving.

There are many other possible explanations for the slight 5-13% decrease in left-right head turns at intersections:

1. Drivers felt duty-bound to focus their gaze more on the pace vehicle than to scan for potential hazards that were never identified, in order to follow their explicit instructions and training to maintain a constant distance to the pace vehicle while performing the tasks.
2. Once the lead vehicle accelerated after stopping or slowing down at the intersection, subject drivers took that as indicating it was safe for them to proceed through the intersection, without the need to scan the intersection.
3. No actual hazard may have been present at any of the pre-defined locations during the experiment.
4. Even if an actual hazard were present, the Utah report does not mention if there were more left-right head turns during a secondary task than during no secondary task.
5. The Utah report does not mention whether there were more actual hazards during task conditions than during the "Single" baseline condition, or whether different task conditions had different numbers of actual hazards.
6. No instructions were given to the drivers regarding the need to search for hazards, and drivers were scored only on the "potential" hazards pre-identified by the experimenters, but never revealed to the subjects.
7. The current study used Google Earth to inspect the actual route used for the Utah on-road study [1, Appendix D] from a ground-level view, revealing that almost all intersections

were free from physical obstruction of sight lines. Thus, when subject drivers were approaching an intersection, they could have seen from a distance that there were no pedestrians or cross-traffic present or approaching those intersections. Hence, there could be a reduced need for left-right head turns during brief stops at such intersections.

8. In any on-road driving study, university Internal Review Board protections ensure that drivers are protected from the conditions of a crash or near-crash, and generally require that drivers be advised that they are so protected. Hence, drivers participate with awareness that their safety is provided for at all times. Indeed, two experimenters were always the vehicle with the subject driver, and the front seat experimenter "notified the driver of any potential roadway hazards" [1, p. 22]. Because the driver knew they would be told of any potential hazards by the experimenter, then it is plausible that the driver would give a lower priority to head turns than any or all of the following tasks: a) following the lead vehicle, b) braking, c) steering, d) responding to the head-mounted red-green detection response task, e) performing the auditory-vocal tasks.
9. No analysis was made of actual glances, only head turns, making it misleading to label the head turn metric as "glances" as was done in the Utah study [1, Figure 26]. Some drivers may have chosen to make lateral glances with their eyes after a partial head turn, instead of making a full head turn. Such glances would not have been detected by the Utah methods, and so such trials would be incorrectly scored as a failure to turn the head.
10. It is not stated how many degrees of neck rotation were considered a head turn as opposed to no head turn. Because the video analysts who made judgments of the head turn amounts could see and hear what the task condition was, there is thus possible observer bias - e.g., the head turns of subjects performing tasks with assumed higher cognitive load could have been scored more strictly, resulting in fewer tabulated head turns.

1.3.5 Other Self-Regulatory Variables Not Mentioned in the Utah Study

A survey study [29] found that the majority of drivers "reported modifying their driving when engaged in other tasks. Most commonly, drivers reported reducing speed, pulling over to side of road, increasing following distance, and stopping their vehicle." The survey also found that "In order to manage their exposure to risk, subjects reported that they will avoid engaging in distracting activities in a range of situations, including poor weather conditions, when driving at night, in heavy traffic, in unfamiliar areas, on winding or curved roads, and around schools." Specific additional self-regulatory variables not reported in the Utah study, but were likely present there, are discussed next.