1. INTRODUCTION

1.1. GOALS

• Use principal component analysis (PCA) to analyze the Crash Avoidance Metrics Partnership Driver Workload Metrics (CAMP-DWM) data to reveal the underlying dimensions of driver performance.

• Examine how these dimensions are associated with key variables used to differentiate secondary visual-manual tasks while driving.

• Show how a two-dimensional approach resolves the apparent paradox of low workload tasks with poor event detection and response, and high workload tasks with good event detection and response.

• Show how these findings clarify the mechanisms underlying relative crash risk for secondary tasks, emphasizing that long single glance durations may be key for missed events that contribute to crashes, more so than eyes-off-road time or number of glances.

1.2. PREVIOUS PRINCIPAL COMPONENT ANALYSES OF DRIVER PERFORMANCE

1.2.1. Young and Angell (2003)

Few experimental studies of driver distraction have been conducted in a moving vehicle, and even fewer use both driver and vehicle performance variables during secondary tasks. One of the first studies to accomplish both goals was conducted on the Smart Road at Virginia Tech Transportation Institute (VTTI) by Angell et al. [1] and analyzed by Young and Angell [2] using principal component analysis (PCA). PCA is a mathematical method for reducing a large number of intercorrelated variables to a smaller subset that can explain most of the variance in a larger set of data ([3], [4], Young [5], Angell et al. [7, Appendix T]).

Young and Angell [2] used PCA to analyze 15 driver performance variables for 79 visual-manual tasks (see Angell et al. [1]). Young and Angell [2, Table 2] found that all 15 driver and vehicle performance variables were correlated in a positive and statistically significant manner with each other. To reduce the interdependencies and simplify the interpretation of the data, the 15 variables were reduced to two principal components, which accounted for 78% of the variance in all 79 tasks across all 15 variables.
1. The first principal component (P1) was loaded with a positive sign for all 15 driver performance variables collected, and accounted for 61% of the total variation of all variables across all tasks. That is, the original profile of the fifteen variables for a task could be reproduced with knowledge of only a single number - the score of that task on the first principal component. PC1 loaded most strongly with the standard driver workload variables of task time, eyes-off-road time, number of glances, lane deviations, speed deviations, percent unsuccessful task completions, situation unawareness, and subjective workload. It loaded weakly with the event detection variables of front and side event miss rates and response times.

2. The second principal component (P2) loaded with a positive sign on event detection and response variables, but a negative sign on workload variables. It identified tasks making drivers less attentive to outside events than expected, given a task's low workload, and accounted for 17% of the remaining total variation. Tasks scoring high on P2 had poor event detection (long response times and high miss rates to hood and side events). However, these tasks were also short duration with low subjective workload, few eye glances and low eyes-off-road time, high subjective situation awareness, and few lane and speed deviations. However, these tasks had relatively high miss rates and long reaction times to a central LED light placed on the hood of the vehicle in front of the driver, as well as a peripheral LED placed on the left outside rear view mirror. These event lights were intended as a surrogate for real-world events such as brake lights turning on in a forward vehicle, or observing an object or event in the left outside rear-view mirror.

The variable loadings of the first and second components are plotted against each other in Fig. 1. Clearly, the workload variables are grouped in the lower right (high positive loadings on P1). On the other hand, the event detection and response variables are grouped in the upper left (high positive loadings on P2). Fig. 1 shows that the mean single glance duration variable “glancedur” is positioned in between the workload variables (in the lower right) and the event detection and response variables (in the upper left). It was (incorrectly) assumed in the Young and Angell [2] study that glance duration must be inherently associated with the driver workload variable of eyes-off-road time, because eyes-off-road time is the product of that variable and the number of glances.

1.2.2. CAMP-DWM Study

The Crash Avoidance Metrics Partnership Driver Workload Metrics (CAMP-DWM) study is described by Angell et al. [6, 7]. The CAMP-DWM partnership was a consortium that included GM, Ford, Toyota, Nissan, and the Intelligent Transportation Systems Joint Program Office, part of the U.S. Department of Transportation. The project lasted four years, from April 2001 to March 2005. Its major goal was to develop driver performance metrics and test procedures to assess how the workload associated with using an in-vehicle system might degrade or interfere with driving performance. This goal was achieved primarily by detailed analysis of individual driver performance variables.

The CAMP-DWM study attempted to use PCA [7, Appendix T] to reduce their 106 driver performance variables to simpler components. However, the PCA methods they used did not converge on a meaningful answer because of three major technical problems: 1) Ill-constraint. There were only 23 tasks for 106 measures, so the data could not be run at the task level; 2) Mixed task types. The auditory-vocal tasks had a fixed duration of 2 minutes, the visual-manual tasks were freely run and had so varying duration; and 3) Missing data. Data at the subject level were missing because of unreduced eye glance data, no glances for auditory-vocal tasks or to some locations for visual-manual tasks, and 9% missing response time data for the FVTS task because of an equipment malfunction. In Section 2.2, we report a new analysis designed to fix or avoid these problems. Note that the unsuccessful PCA attempt by the CAMP-DWM
investigators does not mean the many results, insights, or accomplishments of that study should be construed adversely.

2. METHODS

2.1. CAMP-DWM STUDY METHODS

2.1.1. Event Stimuli

The CAMP-DWM study used actual road events (Fig. 1) rather than lights on the vehicle as in Angell et al. [1], to measure event detection during secondary tasks. The CAMP-DWM study used both open road and track venues, but the track venue had the most tasks, and was most closely related to the closed-road Angell et al. [1] venue, so just the CAMP-DWM track data are analyzed here.

The event detection and response scenarios in the CAMP-DWM study are illustrated in Fig. 2. From left to right these were:

1. Turn Signal Illumination. The following-vehicle turn signal (FVTS) was activated (left-most vehicle in Fig. 2). This event was usually viewed in the left driver side mirror of the center vehicle in Fig. 1. The FVTS illuminated for a fixed 2.5 second duration, and drivers in the center vehicle responded by pressing a button with the left index finger.

2. CHMSL Illumination. The center high-mounted stoplight (CHMSL) was activated, on the top rear of the forward vehicle (right-most vehicle in Fig. 2). The CHMSL would turn on for a duration equal to the instantaneous time headway at stimulus onset (i.e., inter-vehicle distance divided by subject vehicle speed at onset) [6, pp. 2-4 to 2-5]. Drivers in the center vehicle also responded by pressing a button on the left index finger.

3. Deceleration of Lead Vehicle. The lead vehicle decelerated (LVD) (right-most vehicle in Fig. 2) without the brake lights turning on. Drivers in the center vehicle responded by gently tapping on the brake pedal.

Note that the CAMP-DWM event stimuli were entirely different from the event stimuli used by Angell et al. [1], which were LED lights mounted on the front hood and left side mirror of the driver's vehicle. Another difference is that in the CAMP-DWM study on the track, only one visual event was scheduled per task trial. Angell et al. [1] had many events per task trial, presented every 3 to 6 seconds. Stimuli were thus less expected in the CAMP-DWM study than in the Angell et al. [1] study, which some consider more realistic from a driver's perspective. However, more events give rise to a more accurate estimate of miss rates and response times. The CAMP-DWM study also defined a miss when the subject failed to respond by the end of the task trial, compared with 2.5 sec in the Angell et al. [1] study.

2.1.2. Response Time Measures

As mentioned in Section 2.1.1, the driver response for the LVD event was a quick tap on the brake. The response for the FVTS and CHMSL events was the press of a finger switch. A response time (RT) was counted if it occurred at any time up until the end of the task time. An event was only counted as a miss (and the response time not recorded) if the event was not responded to by the end of the task. RTs could thus be as long as the task time (many seconds). For example, the CAMP-DWM report [7, Table Q-38] shows that the mean response time for the “route tracing” task was 6.67 seconds, with a maximum response time of 15.57 seconds. The RTs in Angell et al. [1] could not be longer than 2.5 seconds.

2.1.3. Driver Performance Variables and Tasks

The test track was the most similar venue to the closed road venue in the Young and Angell [2] study. The CAMP-DWM study had 106 driver performance variables and 13 visual-manual tasks in the test track venue. There were 69 subjects tested using instrumented vehicles, for which 42 had eye data scored. Each subject did all 13 visual-manual tasks (auditory-vocal tasks were also tested, but they are not analyzed here). The mean data for many of the 106 variables across all the subjects for all the tasks was provided in Appendix Q [7]. The number of variables analyzed by PCA cannot be larger than the number of tasks or an error occurs because the data are ill constrained. A maximum of 13 out of the 106 variables could therefore be analyzed at any one time, using the mean task data across subjects. Ten variables were found to have a reasonably close match to the Young and Angell [2] variables, and those were the ten selected for PCA analysis, as described in Section 2.2.

2.2. CURRENT PRINCIPAL COMPONENT METHODS

A new PCA analysis of the CAMP-DWM study data seemed warranted. We employed the PCA methods of Young and Angell [2], which avoided or fixed the technical problems in Section 1.2.2. Applying PCA to the CAMP-DWM data could accomplish four things: 1) validate the original Young and Angell [2] approach and components; 2)
cross-validate the CAMP-DWM methods and conclusions; 3) allow for individual tasks to be scored by only two variables instead of dozens, permitting an easier understanding of their meaning, and simplifying the setting of driver performance criteria; and 4) provide a foundation for a later more complete analysis of the full 106 variable CAMP-DWM data set.

2.2.1. Matching Variables between Studies

Only track data (rather than open road or lab data) were analyzed from the CAMP-DWM study, because more tasks were run on the track than on the road, and the track venue had been used by Angell et al. [1], and analyzed by Young and Angell [2]. The ten CAMP-DWM variables that most closely matched the variables in the Young and Angell [2] study were chosen for the PCA.

Some variables were dropped from the PCA for the following reasons. Five of the 15 variables from the Young and Angell [2] study were dropped because: 1) the three variables of subjective workload, subjective situation awareness, and percent unsuccessful task completions did not have a direct match to any variable in the test track venue in the CAMP-DWM study; and 2) the two variables of combined miss rate and combined response time were a linear combination of the side and front events, potentially causing an ill-conditioned PCA. Also, the “following vehicle turn signal” (FVTS) variables in the CAMP-DWM study [6] were dropped because they had the weakest statistical power of the three road events because: 1) the FTVS variable had unreliable RTs because 43% or more of subjects missed the FVTS event even when just driving with no secondary task and so had no RT; 2) the FVTS miss rate had little ability to discriminate tasks, in part because it approached a ceiling of 100% misses for some subjects; and 3) the FTVS event had more artifacts in its RTs than the other two events introduced by an equipment failure late in the CAMP-DWM study [7, page T-20]. Therefore, the current analysis was restricted to the CHMSL and lead vehicle deceleration (LVD) events. Dropping the peripheral FVTS event did not lose explanatory power in a two-dimensional PCA, because the Young and Angell [2] study showed negligible difference in the loadings of the central and peripheral events on the first two principal components. Little difference between the response times and miss rates for a peripheral and central light position has been found in other studies ([8],[9],[10]; see Young [13] for review).

The final ten variables selected for the current PCA are given in Table 1. Column 1 gives the shorthand names in the current study. Column 2 gives their names, and Column 3 the table number, in the CAMP-DWM study (Angell et al. [2, Appendix Q]). The last column is the variable name in Young and Angell [2]. Variables 1-5 are conventional driver workload variables, and variables 7-10 are event detection and response variables. Variable 6 “GlanceDur” was investigated in this study as to whether it was best associated with the event detection and response variables or the workload variables.

### Table 1. Variable names in current study derived from CAMP DWM study [6,7], and closest equivalent variables for Young and Angell [2]. All variables are task-related; that is, they were recorded only while the driver was doing the secondary task on the road.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 TaskTime</td>
<td>Task duration</td>
<td>Q-34</td>
<td>Task duration</td>
</tr>
<tr>
<td>2 EORT</td>
<td>Mean duration of glances</td>
<td>Q-45</td>
<td>Eyes-Off-Road Time</td>
</tr>
<tr>
<td>3 Glances</td>
<td>Mean # glances</td>
<td>Q-44</td>
<td># Glances</td>
</tr>
<tr>
<td>4 SDLP</td>
<td>Standard deviation of lane position</td>
<td>Q-35</td>
<td># Lane deviations</td>
</tr>
<tr>
<td>5 SpeedDev</td>
<td>Speed difference</td>
<td>Q-36</td>
<td># Speed deviations</td>
</tr>
<tr>
<td>6 GlanceDur</td>
<td>Mean single glance duration</td>
<td>Q-46</td>
<td>Mean single glance duration</td>
</tr>
<tr>
<td>7 CHMSL%</td>
<td>% misses of CHMSL event</td>
<td>Q-39</td>
<td>% Misses of hood event</td>
</tr>
<tr>
<td>8 LVD%</td>
<td>% misses of LVD event</td>
<td>----</td>
<td>% Misses of side event</td>
</tr>
<tr>
<td>9 CHMSL_RT</td>
<td>Response time to CHMSL</td>
<td>Q-38</td>
<td>RT to hood event</td>
</tr>
<tr>
<td>10 LVD_RT</td>
<td>RT to LVD</td>
<td>Q-37</td>
<td>RT to side event</td>
</tr>
</tbody>
</table>

Notes
(a) all variables tabulated only during task
(b) glances to displays or controls for task
(c) continuous measure of lane keeping during task
(d) (max - min) speed during task [6, Fig. 4-74]
(e) “center high mount stop lamp” event, [6, Fig. 3-1]
(f) “lead vehicle deceleration” event [6, Fig. 3-17]
(g) response time to CHMSL event
(h) response time to LVD event [6, Fig. 3-18]
(i) data not in Appendix Q; see [6, Fig. 3-17]
(j) # times tire touched lane marker during task
(k) # speed deviations < 35 or > 45 mph during tasks
(l) “hood event” [2] is activation of red LED light on hood of vehicle
(m) “side event” [2] is an LED light above the left outside mirror

The mean task data for the CAMP-DWM study are given in Appendix Q of Angell et al. [7]. Variable 8 (LVD%) did not have task means published in Appendix Q, but these were plotted in Fig. 3-17 of Angell et al. [6]. The numerical values for that variable were requested and received from one of the collaborating partners in the CAMP-DWM study.

2.2.2. Statistical Methods

PCA gives both a classification of driver distraction variables into the major few components (dimensions) that affect driving performance, as well as scores for individual tasks on those dimensions. The principal components and the
scores of the tasks on those components were calculated for the chosen subset of the CAMP-DWM variables. The “MiniTab 16” [11] statistical package was used for all calculations and plots.

We applied the methods of Young and Angell [2] to the CAMP-DWM data set in an attempt to avoid the problems described in Section 1.2.2 with the CAMP-DWM PCA methods [6, Appendix T]. Three specific changes from the CAMP-DWM PCA were made: 1) analyzing task means across subjects, rather than analyzing subject data at an individual level, thereby completely avoiding missing data; 2) restricting the analysis to the ten variables that most closely matched those in Young and Angell [2]), and 3) restricting the data set to just visual-manual tasks, avoiding auditory-vocal tasks with no glances to a device and hence with missing glance durations.

3. RESULTS

3.1. TASK MEANS FOR TEN KEY DRIVER PERFORMANCE VARIABLES

Table 2 gives the means across subjects of the tasks for the selected variables in Table 1 (data from Appendix Q from the CAMP-DWM study [7]).

Table 2. Task means for CAMP-DWM data [7], Appendix Q. See Table 1 for description of the ten variables.

<table>
<thead>
<tr>
<th>TaskName</th>
<th>Task</th>
<th>1 Task</th>
<th>2 EORT</th>
<th>3 Glances</th>
<th>4 SDLP</th>
<th>5 SpeedDev</th>
<th>6 GlanceDur</th>
<th>7 CHMSL %</th>
<th>8 LVD</th>
<th>9 CHMSL_RT</th>
<th>10 LVD_RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ReadHard</td>
<td>30.9</td>
<td>15.3</td>
<td>13.7</td>
<td>0.64</td>
<td>5.73</td>
<td>1.04</td>
<td>18.8</td>
<td>26.9</td>
<td>2.26</td>
<td>5.76</td>
<td></td>
</tr>
<tr>
<td>Cassette</td>
<td>15.9</td>
<td>4.7</td>
<td>9.1</td>
<td>0.63</td>
<td>4.79</td>
<td>1.03</td>
<td>15.8</td>
<td>46.3</td>
<td>2.19</td>
<td>5.49</td>
<td></td>
</tr>
<tr>
<td>CDTtrack7</td>
<td>18.4</td>
<td>9.8</td>
<td>9.3</td>
<td>0.63</td>
<td>4.79</td>
<td>1.03</td>
<td>15.8</td>
<td>46.3</td>
<td>2.19</td>
<td>5.49</td>
<td></td>
</tr>
<tr>
<td>ManualDia</td>
<td>24.4</td>
<td>11.0</td>
<td>11.3</td>
<td>0.68</td>
<td>5.67</td>
<td>0.97</td>
<td>30.2</td>
<td>24.5</td>
<td>2.30</td>
<td>5.94</td>
<td></td>
</tr>
<tr>
<td>Coins</td>
<td>17.6</td>
<td>6.7</td>
<td>7.3</td>
<td>0.65</td>
<td>4.52</td>
<td>0.90</td>
<td>26.4</td>
<td>51.2</td>
<td>2.39</td>
<td>5.94</td>
<td></td>
</tr>
<tr>
<td>ReadEasy</td>
<td>21.3</td>
<td>10.7</td>
<td>10.1</td>
<td>0.62</td>
<td>5.13</td>
<td>1.03</td>
<td>30.2</td>
<td>48.1</td>
<td>2.19</td>
<td>5.78</td>
<td></td>
</tr>
<tr>
<td>MapEasy</td>
<td>17.4</td>
<td>6.6</td>
<td>6.7</td>
<td>0.65</td>
<td>4.53</td>
<td>0.96</td>
<td>33.3</td>
<td>51.8</td>
<td>2.24</td>
<td>5.46</td>
<td></td>
</tr>
<tr>
<td>DestEntry</td>
<td>104.5</td>
<td>57.6</td>
<td>50.3</td>
<td>0.84</td>
<td>8.47</td>
<td>1.15</td>
<td>27.5</td>
<td>31.4</td>
<td>2.72</td>
<td>6.62</td>
<td></td>
</tr>
<tr>
<td>RadioHard</td>
<td>16.7</td>
<td>9.0</td>
<td>7.7</td>
<td>0.59</td>
<td>4.34</td>
<td>1.16</td>
<td>34.2</td>
<td>38.5</td>
<td>2.77</td>
<td>5.42</td>
<td></td>
</tr>
<tr>
<td>MapHard</td>
<td>25.2</td>
<td>12.7</td>
<td>10.7</td>
<td>0.74</td>
<td>5.63</td>
<td>1.10</td>
<td>32.2</td>
<td>46.5</td>
<td>2.61</td>
<td>6.83</td>
<td></td>
</tr>
<tr>
<td>RouteTrace</td>
<td>26.9</td>
<td>12.9</td>
<td>11.4</td>
<td>0.82</td>
<td>6.14</td>
<td>1.19</td>
<td>33.1</td>
<td>41.8</td>
<td>2.96</td>
<td>6.67</td>
<td></td>
</tr>
<tr>
<td>RadioEasy</td>
<td>11.4</td>
<td>5.7</td>
<td>4.6</td>
<td>0.52</td>
<td>3.73</td>
<td>1.28</td>
<td>33.3</td>
<td>72.7</td>
<td>2.38</td>
<td>5.93</td>
<td></td>
</tr>
<tr>
<td>HVAC</td>
<td>11.4</td>
<td>6.3</td>
<td>5.6</td>
<td>0.55</td>
<td>3.84</td>
<td>1.15</td>
<td>34.8</td>
<td>72.3</td>
<td>2.69</td>
<td>6.24</td>
<td></td>
</tr>
</tbody>
</table>

3.2. CORRELATION BETWEEN DRIVER PERFORMANCE VARIABLES

Table 3 gives the correlations between the 10 driver performance variables for the 13 tasks in Table 2. Table 3 shows that there are strong positive inter-correlations between all five driver workload variables in the upper left quadrant (TaskTime, EORT, Glances, SDLP, and SpeedDev). Many of the event detection variables are also positively intercorrelated among themselves (lower right quadrant), but not as strongly as are the driver workload variables. The two groups of variables (1-5 vs. 6-10) have three statistically significant correlations between them, two negative. Note that variable 6 “GlanceDur” (mean single glance duration to the task) is correlated with the event detection and response variables rather than the workload variables: It correlates 0.606 and 0.573 with variables 9 and 10, the response times for the two forward events, CMHSL and lead vehicle deceleration. Variable 6 (GlanceDur) is not correlated in a statistically significant manner with any of the driver workload variables 1-5.

Table 3. Correlation coefficients between variables in Table 2. Highlighted r values are statistically significant at p < 0.05 for n = 13.

<table>
<thead>
<tr>
<th>Task</th>
<th>Task</th>
<th>1 Task</th>
<th>2 EORT</th>
<th>3 Glances</th>
<th>4 SDLP</th>
<th>5 SpeedDev</th>
<th>6 GlanceDur</th>
<th>7 CHMSL %</th>
<th>8 LVD</th>
<th>9 CHMSL_RT</th>
<th>10 LVD_RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ReadHard</td>
<td>1 TaskTime</td>
<td>0.996</td>
<td>0.998</td>
<td>0.711</td>
<td>0.914</td>
<td>0.180</td>
<td>-0.157</td>
<td>-0.469</td>
<td>0.302</td>
<td>0.428</td>
<td></td>
</tr>
<tr>
<td>Cassette</td>
<td>0.996</td>
<td>1</td>
<td>0.998</td>
<td>0.684</td>
<td>0.901</td>
<td>0.248</td>
<td>-0.147</td>
<td>-0.434</td>
<td>0.326</td>
<td>0.457</td>
<td></td>
</tr>
<tr>
<td>CDTtrack7</td>
<td>0.998</td>
<td>0.998</td>
<td>1</td>
<td>0.690</td>
<td>0.906</td>
<td>0.198</td>
<td>-0.165</td>
<td>-0.459</td>
<td>0.298</td>
<td>0.432</td>
<td></td>
</tr>
<tr>
<td>ManualDia</td>
<td>0.711</td>
<td>0.684</td>
<td>0.690</td>
<td>1</td>
<td>0.887</td>
<td>0.046</td>
<td>-0.049</td>
<td>-0.574</td>
<td>0.461</td>
<td>0.599</td>
<td></td>
</tr>
<tr>
<td>Coins</td>
<td>0.914</td>
<td>0.901</td>
<td>0.906</td>
<td>0.887</td>
<td>1</td>
<td>0.119</td>
<td>-0.199</td>
<td>-0.665</td>
<td>0.314</td>
<td>0.539</td>
<td></td>
</tr>
<tr>
<td>ReadEasy</td>
<td>0.180</td>
<td>0.248</td>
<td>0.198</td>
<td>0.046</td>
<td>0.119</td>
<td>1</td>
<td>0.380</td>
<td>0.313</td>
<td>0.606</td>
<td>0.573</td>
<td></td>
</tr>
<tr>
<td>MapEasy</td>
<td>-0.157</td>
<td>-0.147</td>
<td>-0.165</td>
<td>-0.049</td>
<td>-0.199</td>
<td>0.380</td>
<td>1</td>
<td>0.401</td>
<td>0.513</td>
<td>0.273</td>
<td></td>
</tr>
<tr>
<td>DestEntry</td>
<td>-0.469</td>
<td>-0.434</td>
<td>-0.459</td>
<td>-0.574</td>
<td>-0.665</td>
<td>0.313</td>
<td>0.401</td>
<td>1</td>
<td>0.022</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td>RadioHard</td>
<td>0.302</td>
<td>0.326</td>
<td>0.298</td>
<td>0.461</td>
<td>0.314</td>
<td>0.606</td>
<td>0.513</td>
<td>0.022</td>
<td>1</td>
<td>0.596</td>
<td></td>
</tr>
<tr>
<td>MapHard</td>
<td>0.428</td>
<td>0.457</td>
<td>0.432</td>
<td>0.599</td>
<td>0.539</td>
<td>0.573</td>
<td>0.723</td>
<td>0.010</td>
<td>0.596</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

3.3. PRINCIPAL COMPONENTS OF DRIVER PERFORMANCE

Table 4 and Fig. 3 show the loading of the variables on the major two principal components for the correlation matrix in Table 3. The total variance explained by P1 and P2 combined is 78%, with 22% residual variance. This high value is consistent with the 78% of variance explained by the first two components in the 15 variable dataset and 79 visual-manual tasks in Young and Angell [2]. The residual variance in the remaining components in the current study was minor, and contained no recognizable pattern, indicating they were likely statistical noise. The component loadings P1 and P2 are orthogonal, meaning they have zero correlation.

Fig. 3 (top panel) illustrates that P1 is mostly positive. The P1 component alone explains 53% of the total variance in the data sample. That is, instead of needing ten variables for every task, the single task score on P1 can reproduce most of the variance across all ten variables in Table 2 for that task. This 53% is comparable to the 61% explained by P1 in the Young and Angell [2] study. Fig. 3 (bottom panel) shows that the second component is a contrast between 1) event variables and single glance duration (right-most positive red bars); vs. 2) the driver workload variables (left-most negative
blue bars). The P2 component explains 25% of the remaining variance in the CAMP-DWM variable set examined, about the same as the 17% of variance explained by the P2 component in Young and Angell [2].

Table 4. Loadings of variables on the major two principal components for the correlation data in Table 3.

<table>
<thead>
<tr>
<th>Variable</th>
<th>P1</th>
<th>P2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaskTime</td>
<td>0.40888</td>
<td>-0.08138</td>
</tr>
<tr>
<td>EORT</td>
<td>0.40827</td>
<td>-0.05076</td>
</tr>
<tr>
<td>Glances</td>
<td>0.40705</td>
<td>-0.07730</td>
</tr>
<tr>
<td>SDLP</td>
<td>0.37305</td>
<td>-0.02450</td>
</tr>
<tr>
<td>SpeedDev</td>
<td>0.42068</td>
<td>0.11141</td>
</tr>
<tr>
<td>GlncDur</td>
<td>0.11795</td>
<td>0.49790</td>
</tr>
<tr>
<td>CHMSL%</td>
<td>-0.03773</td>
<td>0.49566</td>
</tr>
<tr>
<td>LVD%</td>
<td>-0.23503</td>
<td>0.37567</td>
</tr>
<tr>
<td>CHMSL_RT</td>
<td>0.20457</td>
<td>0.45521</td>
</tr>
<tr>
<td>LVD_RT</td>
<td>0.26809</td>
<td>0.36040</td>
</tr>
</tbody>
</table>

Figure 3. Loadings of the original variables on the first two components for the visual-manual tasks in the CAMP-DWM track study [6]. This is a plot of the variable loadings in Table 4.

Figure 4 plots the P2 loadings vs. the P1 loadings. The event-related variables load high on P2 and low on P1 (red squares, upper left). The driver workload variables load high on P1 and low on P2 (black circles, lower right). The glance duration variable GlnceDur is clearly grouped with the event detection and response variables in the upper left rather than the driver workload variables in the lower right.

3.4. TASK SCORES ON PRINCIPAL COMPONENTS

The scores on just the two components P1 and P2 can reproduce 78% of the total variation in the ten original variables for all 13 task means in Table 2. Table 5 gives the weights or scores on the tasks that, when cross-multiplied with the component loadings and summed, can reproduce the original profiles of the task means across all ten variables to a high accuracy.

Table 5. Task scores (weights) on the principal component loadings that reproduce the original data set using only two scores S1 and S2 instead of 10 variables.

<table>
<thead>
<tr>
<th>Task</th>
<th>S1</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task</td>
<td>0.4630</td>
<td>-1.9598</td>
</tr>
<tr>
<td>ReadHard</td>
<td>-1.9544</td>
<td>-1.9002</td>
</tr>
<tr>
<td>Cassette</td>
<td>-0.9033</td>
<td>-1.8605</td>
</tr>
<tr>
<td>CDTrack7</td>
<td>0.2934</td>
<td>-1.1174</td>
</tr>
<tr>
<td>ManualDial</td>
<td>-0.9893</td>
<td>-0.7033</td>
</tr>
<tr>
<td>Cords</td>
<td>0.7095</td>
<td>-0.4819</td>
</tr>
<tr>
<td>MapEasy</td>
<td>1.3465</td>
<td>-0.4110</td>
</tr>
<tr>
<td>DestEntry</td>
<td>6.6562</td>
<td>-0.2750</td>
</tr>
<tr>
<td>RadioHard</td>
<td>-0.7470</td>
<td>1.0628</td>
</tr>
<tr>
<td>MapHard</td>
<td>0.9638</td>
<td>1.2055</td>
</tr>
<tr>
<td>RouteTrace</td>
<td>1.8757</td>
<td>1.9701</td>
</tr>
<tr>
<td>RadioEasy</td>
<td>2.0335</td>
<td>2.0871</td>
</tr>
<tr>
<td>HVAC</td>
<td>-1.5687</td>
<td>2.3835</td>
</tr>
</tbody>
</table>
The task scores are orthogonal with each other (zero correlation between S1 and S2), just as were the component loadings in Table 4. Having both loadings and scores orthogonal is a unique feature of principal components, as opposed to any other linear combination of the variables. Any rotation of the axes loses orthogonality in either the loadings or score vectors [3]. There is therefore a unique set of principal components for any given set of variables.

The scores in Table 5 are graphed in Fig. 5. The top panel shows that the “DestEntry” task (black bar) had the highest score on the P1 component. The other tasks are all near zero. However, the driver workload component P1 is of no value in discriminating these other tasks. P1 is a one-dimensional variable (even though it loads with five driver workload variables, these are all highly intercorrelated and contribute no individual information). Analogously, a person with monochromatic color vision can only see things in shades of gray, the “first dimension” of color vision. Likewise, the first dimension of driver performance can only see tasks in “shades of gray” as par the shading of the tasks in the top panel in Fig. 5.

These bars are colored red to indicate they are associated with event variables that have poor event detection and response to a red CHMSL brake light or deceleration of the lead vehicle, despite their having relatively low workload scores on P1.

2. The second group is the five tasks that score intermediate on P2 (colored gray in the middle of the bottom panel of Fig. 5). P2 cannot “see” any difference between these tasks. This group includes DestEntry, which has a high P1 score, but which P2 sees as little different from zero in its event detection and response score.

3. The third group is the three tasks that score negatively on P2 (colored blue to the left of the bottom panel of Fig. 5). These tasks have better event detection and response, and shorter single glance durations, than the other tasks.

The task scores on P2 are shown in the bottom panel of Fig. 5. These tasks had low workload on the P1 component, and could only be represented as “shades of gray” in that dimension, continuing the analogy to color vision. However, the P2 dimension now successfully separates these tasks into three groups, adding a new dimension of “color” to the tasks not previously seen:

1. The first group is the five tasks that score positively on P2 (colored red in the right of the bottom panel of Fig. 5).

2. The second group is the five tasks that score intermediate on P2 (colored gray in the middle of the bottom panel of Fig. 5).

3. The third group is the three tasks that score negatively on P2 (colored blue to the left of the bottom panel of Fig. 5).

Fig. 6 plots the P1 and P2 task scores against each other to show the positioning of each task in the two dimensional space spanned by the first two principal components. Note that the P2 dimension separates the lower-workload tasks which were all seen as the same by P1 into three groups (red squares, grey triangles, blue diamonds; the colors for the tasks are the same as in the lower panel of Fig. 5). The P1 dimension separates out only one of the tasks (DestEntry), seeing all others as the same. Instead, the P2 dimension clearly separates the tasks into three groups that vary in their ability to affect event detection and response, even though all have relatively low driver workload. A new group of tasks is identified as those with poor event detection and response,
even though they have relatively low driver workload (red squares in upper left of Fig. 6).

3.5. CONTROL FOR TASK TIME

Note in Table 3 that the four driver workload variables 2-5 are all highly positively correlated with TaskTime (variable 1). However, they were not correlated in a statistically significant manner with driver event detection and response (variables 6-10). Perhaps this differential result was an artifact of task time. For example, increased task time would plausibly increase the opportunity for glances to a device, or for increased lane variability and speed deviations. However, task time might not affect an event variable because there was only one event per trial in the CAMP-DWM track venue, no matter what the duration of the task.

To control for this possible artifact, the driver workload variables 1-5 were divided by task time to yield rates or proportions, as is commonly done in many driver performance studies (e.g., conversion of an absolute duration into a proportion of time looking at a device). A conversion to rates or proportions is best done at the individual trial level, but those data were not available in the published CAMP-DWM data. Therefore, the conversion was done at the task level as the best available opportunity. The PCA now had to discard task time as a variable because it was now normalized to a constant value “1” and PCA will fail with a constant variable. PCA of the remaining nine variables showed some slight variations in the loadings, but the final locations of the task scores in the 2-D space formed by the P1 and P2 components with transformed variables were not different from those shown in Fig. 6. That is, the pattern of the task scores in Fig. 6 was robust, after controlling for task time.

4. DISCUSSION

The principal components of driver performance as revealed by this analysis are similar to those in Young and Angell [2]. The common result from these two different data sets demonstrates that event detection-response and driver workload represent fundamentally different classes of variables for assessing driver performance and distraction. The first two major dimensions underlying these two classes of variables are orthogonal, so variables in one class cannot substitute for those in the other. Orthogonality also means that the two dimensions cannot be validly combined to give rise to a single driver performance metric - the two dimensions are measuring two intrinsically different aspects of driver performance. Hence, it is necessary to measure both event detection-response variables and driver workload variables to have a complete picture of task properties that may adversely affect driver performance and relative crash risk. If both dimensions are not assessed, tasks could be missed that might adversely affect driving safety.

These results were achieved by applying the same methods of principal components analysis (PCA) used by Young and Angell [2] to the CAMP DWM driving performance data [6,7]. Ten variables that most closely matched those in the Young and Angell [2] dataset were chosen from among the 106 variables in the CAMP-DWM study for the track venue. The PCA method reduced the ten CAMP-DWM driver performance variables to two, while still explaining 78% of the variance in the full set of ten variables, the same amount as explained by the first two components in Young and Angell [2]. The remaining 22% of the variance was largely statistical noise. Thus, ten variables could be reduced to two components with a loss of only a small percentage of the total variance in the original variables. These two components represent two underlying fundamental dimensions of driver performance and distraction for visual-manual tasks, which are independent and orthogonal.

The first and most prominent principal component (P1) explained about 53% of the total variance among the ten variables, and was associated with the commonly used driver workload variables of task time, eyes-off-road time, number of glances, standard deviation of lane position, and speed differences. All these variables are highly intercorrelated and all measure the same thing - driver workload - so they can be effectively compressed into a single component. This result confirms the identification of the first component of Young and Angell [2] as “driver workload.” It is also associated with drivers' subjective impression of the amount of workload they undergo as they perform a task [2].

The second principal component (P2) explained about 25% of the remaining total variance among the ten variables examined in the CAMP-DWM study for visual-manual tasks. The P2 component loaded most heavily with event detection and response variables, but also loaded with mean single glance duration. This latter finding is new, an effect that was present but not noted in the original PCA study by Young and Angell [2]. It was mistakenly assumed there that all glance-related variables must necessarily load together on the driver workload component P1, and that none could load on the event detection and response component P2.

The scores of the 13 tasks on these components was then calculated and plotted. Instead of having to interpret ten variables for each task, only two variables have to be interpreted - the scores on the principal components. These scores can be easily viewed and interpreted in a simple two-dimensional plot, instead of having to interpret ten (or more) variables that are highly correlated (collinear), an almost impossible task. A new discovery was made of a class of low-workload tasks that have poor event detection and response.

4.1. APPARENT EVENT DETECTION “PARADOX” IS INSTEAD A NEW DISCOVERY

This study resolves the apparent paradox of why some tasks can be short with little workload (i.e., score low on P1), but still have poor event detection and response (i.e., score high on P2). (These tasks are represented by the red squares...
in the upper left of Fig. 6.) Conventional wisdom is that tasks with low driver workload should not create difficulties in driver performance or distraction. Note that the single glance duration variable loads on P2, not on P1 (Fig. 4). Hence, single glance duration (to a secondary visual-manual task) is associated with event detection and response while the number of glances and eyes-off-road time are not. Hence, low workload tasks can still have poor event detection and response, if they tend to elicit long single glances. The characteristics of tasks that give rise to long single glances are unknown. A preliminary suggestion will be presented in Section 4.3. We now present evidence that such tasks exist, and show how the apparent paradox can be resolved by considering two dimensions rather than one.

The Young and Angell [2] study was the first to report that some short duration tasks could have relatively worse object and event detection and response. This finding was confirmed in the Crash Avoidance Metrics Partnership Driver Workload Metrics (CAMP-DWM) study (Angell et al. [6, p. 8-6]), where it was reported that this result “holds for simple light stimuli presented either alone or while driving in a simulator such as a peripheral detection task (PDT) Alone, or PDT performed with a simulator.” The CAMP-DWM study also showed that this result holds for road signs when compared with a memorized set (Sternberg spatial and visual stimuli). In addition, it holds for CHMSL and FVTS onsets from other vehicles while driving for road and track venues. The CAMP-DWM study [6, p. 8-6] referred to this finding as “paradoxical”: “Paradoxical results emerged for OED [object and event detection] performance. Shorter and ostensibly less distracting tasks were associated with poorer detection performance than longer and ostensibly more distracting tasks.” Various control experiments investigated by the CAMP-DWM investigators showed that the “paradoxical” effect was not an artifact. The finding that some low workload tasks can have worse event detection than high workload tasks held regardless of whether there were multiple or single events per task trial, or whether there was a time threshold for a miss or not.

However, the current study suggests that the event detection “paradox” is not a paradox or artifact at all, but is rather a new discovery about the relationship between event detection and driver workload. The “paradox” is a paradox only if tasks are rated on driver distraction along a single dimension. The apparent paradox is resolved if driver distraction is considered in two dimensions rather than one. In two dimensions, tasks can score low on dimension one, and at the same time score high on dimension two (or vice versa). It is no longer a paradox that some short tasks can simultaneously have low workload and poor event detection. Tasks can independently score high on dimension one (long task times) and have good event detection (negative or neutral scores on dimension 2). Alternatively, short tasks (low on dimension one), can independently have poor event detection (score high on dimension two). It has been recognized for some time that driver distraction is multivariate in nature. The new finding of the current study is to show how multiple variables can be clustered into two groups, when looked at in two orthogonal dimensions. Thus the apparent event detection paradox, instead of being an artifact to be eliminated or controlled in future studies, is instead a confirmation of the second dimension in driver performance as originally proposed by Young and Angell [2]. This orthogonal second dimension indicates that event detection and response during visual-manual tasks is independent of the driver workload of those tasks.

4.2. MEANING OF LONG SINGLE GLANCE DURATIONS

This analysis of the CAMP-DWM data indicates that single glance durations are more closely related to event detection and response than to driver workload, unlike what was assumed in the Young and Angell [2] study (see Section 1.2.1). Glance duration and eyes-off-road-time load separately on orthogonal dimensions, and so provide independent contributions to driver performance. Clearly, mean single glance duration contributes uniquely to event detection and response, independently of driver workload, explaining variance in P2 that is not explainable by P1.

Specifically, the current analysis shows that “mean single glance duration” is of predictive value in estimating event detection and response nonperformance (misses and long reaction times) for visual-manual tasks. Mean single glance duration is the only such glance variable examined here that did so, because eyes-off-road time and number of glances were associated with driver workload variables on the orthogonal dimension P1. This result means that single glance durations make an independent contribution to event detection and response performance while driving, unlike eyes-off-road time or number of glances. Despite the fact that eyes-off-road time is the product of the number of glances and the mean single glance duration, the contribution of eyes-off-road time to the driver workload dimension P1 comes entirely from the number of glances, rather than mean single glance duration.

Why should events be missed or responded to more slowly if single glance durations to the secondary visual-manual task become longer? One answer is easy and plausible - the longer a driver looks off the roadway in a single glance to a task-related device inside the vehicle, then the greater the probability that a safety-related roadway event will occur while the person is looking at the device. Assume a roadway event occurs while the driver is looking at the device. When the driver finally looks back to the roadway, assume the event can still be seen, and a braking response or steering maneuver is undertaken. The response will be delayed by at least the duration of the time between when the event first occurred, and the driver returned their eyes to the roadway. If the driver made only a short duration glance to the secondary task, the delay in response to the road event would not be as large as with a long duration glance.
However, the deeper question is - why does a driver sometimes look to a device for longer than usual? Usually drivers look back to the roadway when performing a task within about 1.5 seconds [28,29,31]. Previous research indicates that drivers tend to avoid eye glance when their tasks are getting past a limit that they feel comfortable with, and will glance back periodically to check the roadway [37]. The CAMP-DWM data and that of Young and Angell [2], both indicate that short dwell times on the device are associated with many glances back and forth to the roadway. These short dwell times are further associated with tasks that have high workload as scored on the P1 dimension. For example, there are many glances associated with alphanumeric data entry for entering destinations in a navigation system, or address book entries, as these and many other studies have shown. When drivers repeatedly glance back and forth to the road and the device, they are subjectively aware of their workload as indicated by the high values on the subjective workload variables (Young and Angell [2]). Their glances to the device tend to be relatively short and they do not miss many roadway events, as shown in both the Young and Angell [2] study and confirmed in the CAMP-DWM study [6]. The repeated number of eye glance off the road “add up” to a long eyes-off-road time, and one might intuitively expect that should increase the number of missed events or increase the response times to those events. These data show however that this intuition is incorrect. Both the Young and Angell [2] study with many events per task trial, and the CAMP-DWM study with only one event per task trial [6], showed no relative decrease in event detection and response performance for long duration tasks. If, however, the task is very short, the driver will not experience subjective feelings of high workload, and may look to the device for longer than usual (see next Section for further discussion).

A contrary argument is that drivers may choose to look away from the road at an in-vehicle task, having already assessed that there was nothing likely to happen on the road ahead for the next several units of time, and so their crash risk would not be increased. In such cases, however, a long single glance duration off the road would still plausibly increase relative crash risk, because a driver can drift out of the lane or even off the road if the eye dwell time off the road is too long.

4.3. SHORT TASKS WITH LONG GLANCES AND POOR EVENT DETECTION

We now further examine the apparent paradox in the previous two sections that some low workload tasks with only one or a few glances off the roadway have worse event detection performance than high workload tasks. Note in Fig. 5 (top panel) that all other tasks scored lower on the P1 dimension than the destination entry task, and they cannot be discriminated from each other just by their scores on the driver workload dimension P1. As pointed out in Section 3.4, these low workload tasks can be independently scored on the P2 dimension (Fig. 5, bottom panel). Among those short tasks, five have poor event detection for external events. However, note that these five tasks also have a long mean single glance duration to the device (Fig. 5, bottom panel, red bars; and Fig. 6 red squares). Fig. 6 shows that the HVAC (Heating-Ventilation-Air-Conditioning) task was the worst task for event detection on the P2 dimension (with 35.4% CHMSL miss rate, 2.69 sec mean RT in Table 2, and the fourth highest mean single glance duration of 1.15 sec). Yet the HVAC task was one of the shortest (11.4 sec) in the CAMP-DWM study. There are at least three possible conjectures for why some low workload tasks can have long single glance durations and poor event detection and response.

4.3.1. “Trying to Finish Task” Conjecture

As drivers get close to finishing a task, they may extend their dwell time so they can finish the task with one long last glance, saving the “extra” workload associated with having to return the eyes to the road and then again back to the device to finish the task. That is, drivers may take extra dwell time in the last glance. For short visual-manual tasks, there are necessarily few glances, so the long final single glance duration will disproportionately increase the overall mean glance duration. (This strategy is a poor one, if it is indeed what is being pursued by drivers, because the current data show that event miss rates and response times substantially increase during tasks when drivers make longer single glances.)

This conjecture, although it seems intuitive at first, has difficulty on close examination of the data across all short tasks. Some short tasks have better than average event detection (such as the “Cassette” and “CDTrack?” tasks seen in Fig. 6 as blue diamonds). These tasks have relatively short task times (15.9 and 18.4 seconds respectively as shown in Table 2) and few glances (5.7 and 9.1), but relatively good event detection and response. On the other hand, compare the “red square” tasks in Fig. 6 that score positively on the P2 scale and have poor event detection and response. Yet these tasks (RouteTrace and MapHard) have longer task times (26.9 and 25.2 seconds respectively), and more glances (11.4 and 10.7). The fact that some relatively short tasks have good event detection and others not indicates that the “trying to finish task” explanation is not sufficient to explain the results across all short tasks. It must be that some short tasks invite “trying to finish” at one time without interruption, and others do not, as described in the next section.

4.3.2. “Interruptability” Conjecture

Another conjecture involves “interruptability” characteristics of tasks, with some tasks being possibly more difficult to interrupt and so having increased dwell time. Monk et al. [12]found that subjects spend more glance time on the secondary task when there are fewer interruptions in that secondary task - that is, when task times are shorter and
there are fewer glances back to the roadway as with longer tasks. Although this interruptability explanation is intriguing, it does not help in determining the particular qualities of the five tasks that score high on P2 (Fig. 5, red squares) that may contribute to drivers making fewer interruptions (glances back to the roadway) with those tasks than with others of equivalent short duration.

4.3.3. “Cognitive Distraction” Conjecture

The study of “cognitive distraction” while driving is still in its infancy. There is wide disagreement in the driving literature field of how to define it [13] and most of the existing variables and models have issues [13]. We conjecture here that the P2 dimension may offer a new way to define “cognitive distraction” in terms of event detection and response performance, at least for visual-manual tasks. Defining cognitive distraction as associated with the P2 dimension gives rise to another conjecture as to why some short duration tasks can have poor event detection and response.

The conjecture is that some short duration visual-manual tasks have the ability to create a “cognitive distraction” state in the driver (however the term is defined [13]). When a person is in a state of cognitively distraction, they are not necessarily aware that they are distracted, a phenomenon known as “attention capture” [14]. As discussed in Section 4.2, if a visual-manual task is long and difficult, such as destination entry, drivers are aware of their subjective feelings of workload, frustration, time pressure, etc. - they realize they are in a distracted or at least uncomfortable subjective state. They experience a “pull” factor to return their eyes to the road, particularly under adverse traffic, weather, or road geometry conditions. Direct evidence is in the data in Fig. 1 - the “workload” data point, referring to drivers’ subjective workload rating of the task right after completion, is tightly clustered with the other driver workload variables. That is, long tasks generate feelings of discomfort, but short tasks do not, in the self-reports of the drivers [2]. Hence, during some short tasks, it is conjectured that drivers may dwell on controls and displays associated with the task for longer than normal because “attention capture” is more possible for those shorter tasks, because of the lack of a subjective driver workload “cue.” Attention capture tends not to occur with longer tasks, because drivers experience an uncomfortable subjective feeling of workload and return their eyes to the roadway more frequently or even shed the task as too difficult. That is, they are more willing to “shed” a high-workload visual-manual secondary task and look back to the road more quickly. They undertake a visual glance strategy comprised of short, quick, and multiple glances between the road and the in-vehicle task. However, a driver during a short visual-manual task would not have the feelings of workload that would prompt them to undertake this visual glance strategy. Consequently, some short tasks may give rise to “attention capture” because the driver would not be aware how long the eyes were off the road in a single glance during a low workload task.

The key underlying issue then becomes why some short tasks exhibit more “cognitive distraction” or “attention capture” than others do. The current analysis shows that some short tasks have better than average event detection (the blue diamonds in the lower left of Fig. 6) and shorter duration glances, and other short tasks have poor event detection (red squares in Fig. 6) and longer duration glances. The design characteristics of the relatively short tasks with poor event detection are unknown. This is a new class of tasks not previously recognized in the driving safety or driver vehicle interface literature. The top three examples found in this study of such low workload in-vehicle visual-manual tasks that score relatively poorly on event detection are given in Appendix A. The commonalities of the tasks described in Appendix A are not immediately apparent from the description of the tasks given. If this commonality could be discovered, it may give a clue to the design properties that affect cognitive distraction and poor event detection and response performance while driving. However, the main purpose of this paper is to show that a variable for this effect is present in the second dimension of driver performance. The opportunity for further research is to discover the characteristics of these and other low workload tasks that give rise to poor event detection. That discovery will enable improved task designs or countermeasures to improve driving safety.

In the current paper, we are not suggesting that cognitive distraction is unrelated to driver distraction. We are also not suggesting that driver workload is unrelated to driver distraction. We are saying that the current analysis shows that driver workload is unrelated to event detection and response for visual-manual tasks, as given by dimension P2. We are also suggesting that cognitive distraction may be more associated with the event-related dimension P2, rather than the driver workload dimension P1. We conjecture that drivers are not cognitively distracted by high workload tasks, because they are aware of their distraction because of the subjective feelings of high workload associated with these visual-manual tasks. However, drivers report low subjective workload for many tasks, and so may be more subject to cognitive distraction associated with those low workload tasks (i.e., they are more prone to “attention capture” by some low workload tasks). Ultimate answers to “cognitive distraction” and its relation to event detection and response and driver workload, will likely require future work exploring how long single glance durations and cognitive distraction are associated with “cognitive attentional networks” [15,16]. These networks have been proposed to explain the brain mechanisms associated with driver distraction measured in lab and on-road driving [17,18,19]. These brain mechanisms also have been identified in brain imaging studies of the effect of cellular conversations while driving on event detection and response [20,21,22].
A complete description of all the tasks is given in Angell [2]. A preliminary analysis of visual and cognitive design factors is given in Appendix A of tasks scoring high on P2 in the CAMP-DWM study. Now that the current study has identified ways to discriminate such tasks from others (using the P2 dimension or the variables that load on it), a more in-depth investigation can be made of what characteristics of these particular low workload tasks engender both poor event detection and long mean single glance durations.

4.4. COGNITIVE DISTRACTION, SINGLE GLANCE DURATIONS, AND CRASH RISK

Although it may seem strange, the data examined here suggest that the duration of a glance is associated with cognitive distraction, while driver workload is not. The data of Young and Angell [2] indicate that the driver is aware of the workload of high workload tasks, and so as discussed in the previous sections, may be “cued” to return eyes to the road, and not be “cognitively distracted” during such tasks. In short, high subjective workload does not imply high cognitive distraction for a visual-manual task.

What then is the relative importance of cognitive distraction to crash risk? Young and Angell [2] found that the driver dimension P2 (associated with event detection and response, and cognitive distraction), accounts for only about one-fourth the variance of the driver dimension P1 (associated with driver workload) in overall driver performance. That is, event detection and response variables account for a smaller percentage of the variance in the overall set of 79 visual-manual secondary tasks investigated by Young and Angell [2] than workload variables. A similar result was found in the in the current study of the CAMP-DWM data.

To the extent that “cognitive distraction,” however defined [13], is associated with event detection and response, the greater strength of workload distraction vs. cognitive distraction that follows is consistent with the conclusions of the CAMP-DWM investigators with the full set of 23 tasks they used (which included audio-vocal as well as visual-manual tasks). They stated that “cognitive distraction” played a much smaller role than “visual distraction” (what we here term “workload” distraction). Angell et al. [6, p. xli] note that, “Cognitive distraction effects are very subtle and are not monolithic. Relative to visual distraction, cognitive distraction accounts for much less of the overall variance in driving performance than visual distraction.” In support of the finding of the relative weakness of cognitive distraction vs. workload distraction, Angell et al. [6, p. 8-27] further state that:

“Identifying any measurable interfering effects of cognitive load on driving performance has proven very difficult, in part due to the fact that it is difficult to devise tasks that impose cognitive-only (or even primarily cognitive) loads on the driver. This is necessary to clearly discern the effects of cognitive load on performance. Usually cognitive load co-occurs with other types of task loading (visual, manual, auditory, vocal) so it is not always clear whether any observed effects on driving performance can be attributed to the cognitive portion of the loading. However, a more important issue is that when the visual and manual loads are eliminated to help isolate the cognitive operations and a task allows the eyes to be forward, on the road and/or scanning the traffic environment, the magnitude of intrusions on event detection, as compared with visual-manual tasks, is smaller and more similar to just driving, as are the magnitude of effects on lane-keeping and speed variability, though some of these effects are discriminable from just driving. A clear interpretation of whether there is an intrusion on driving from cognitive load is thus very difficult.”

We show in this paper that the PCA approach here and in Young and Angell [2] have made it possible for the first time to separate the event-related and workload properties of visual-manual tasks. It is thus possible using these methods to provide a clear interpretation of whether there is an intrusion on driving from “cognitive load” in so far as it is reflected in event detection and response variables, a problem that was called “very difficult” in the last words in the quote.

A new finding of this study is that single glance duration was more closely associated with event detection and response variables than with any of the conventional driver workload variables - task time, lane variability, speed deviations, and so forth. That indicates that single glance durations are conveying specific information about event detection and response capabilities that is not being assessed by conventional driver workload variables.

The importance of single glance duration in the second driver performance dimension (and its possible relation to cognitive distraction) in turn may be useful for improved estimates of the relative crash risk of secondary tasks. The use of single glance duration to a secondary task as a sensitive indicator of crash risk was proposed historically a number of years ago in several independent studies [23,37,28,29,30,31]. Recent re-analysis of the 100-car data is also supportive of this association. A close examination of the glance patterns in the 100-car data confirms that single glance duration provides the most sensitive indicator of crash risk of those examined [24,34]. Liang et al. [24] looked at 24
different algorithms that combined glance duration, glance history (eyes-off-road time), and glance location in various combinations. They found that the variable of single off-road glance duration and not glance history was the best crash predictor. Adding glance history and glance location as additional variables did not improve estimation above glance duration variables. Using a short time window around the time of the precipitating event also led to better performance than a large time window. These investigators concluded, “The distraction level estimated by the algorithms that consider current glance duration provides the most sensitive indicator of crash risk.” The study by Victor and Doza [34] extended the Liang et al. [24] observations by calculating odds ratios for the different glance variables as a function of their duration, in the 100-car crash and event database. The largest odds ratios for glance durations of one second or longer were always largest for the glance that overlapped the time of the precipitating event onset (such as a braking lead vehicle), compared to total glance time, the last single glance, or the glance history. Furthermore, the odds ratios (i.e., relative crash risk) steadily increased from glance durations of 1 -1.5 seconds (2.69, 95% CL 1.77-2.69) to 1.5-2 sec (3.02, 95% CL 1.76-5.20) to more than two seconds (3.61, 95% confidence limit 2.44-5.35). They conclude that, “The 100-car data is really saying that single glances from one second and up at the wrong time (at precipitating event) are significantly risky, not necessarily accumulated glance history” [33]. These latest results are consistent with the effect found in the present data and now evident in the Young and Angell [2] study in hindsight (see Fig. 6). The fact that mean single glance duration is associated with event detection, but eyes-off-road time and number of glances are not, in conjunction with the crash analyses by [24,34], indicates that event detection variables (and long single glance durations) may be more important for estimating relative crash risk of secondary tasks than eyes-off-road time or other commonly used driver workload variables. If true, the question remains open as to why both the Young and Angell [2] and Angell et al. [6] data indicate that the second event detection dimension accounts for a much smaller percentage of the overall variance than does the first driver workload dimension. It is an open research question as to why long single glance durations near the time of the precipitating event may be a much better predictor of relative crash risk than eyes-off-road time accumulated over many glances.

The finding of the importance of mean single glance durations is slightly different but consistent with the results of Horrey and Wickens [25]. These investigators point out that “In general, the unsafe conditions that are likely to produce a motor vehicle crash reside not at the mean of a given distribution (in other words, under typical conditions), but rather in the tails of the distribution.” They show some empirical examples of how using variables that account for the upper end of a histogram distribution (beyond the normal thresholds that are often used to classify driver responses), can be even more predictive than the mean. Therefore, the Horrey and Wickens [25] results do not contradict the findings here that the mean single glance duration can be a useful variable - it contributes to the second dimension of driver performance. The Horrey and Wickens paper suggests that the right tail of the single glance duration distribution may contain even more information that is crash relevant to safety than the mean (that is, a safety-related predictor might not be found just using the mean, but might have been found if the right tail of the distribution had been analyzed). So additional analysis of the RT [26] and miss rate histograms may reveal additional variables that are even more predictive of crash rates. If a result is found using a mean, the Horrey and Wickens conclusion does not negate that finding; they indicate that the predictive abilities might be even greater if the right tail of the distribution were also analyzed. So the current study is consistent with Horrey and Wicken's in recommending that “data mining” of the information in complete histograms of the single glance durations under different cognitive loads induced by secondary tasks is a potentially rich research opportunity for understanding event detection and response performance during driving.

The current study also suggests that single glance duration is a key variable that is emerging with new importance, particularly for its role in cognitive distraction. This is not to say that other glance variables do not also play a key role in distraction. Eyes-off-road time and number of glances, as well as scan patterns in general, undoubtedly play an important role in understanding driver performance when doing secondary tasks. The new finding here is that single glance duration may be a particularly useful surrogate for event detection and cognitive distraction for visual-manual tasks, more so than number of glances or eyes off road time. However, these other glance variables relate to P1 and driver workload. However, the current study does show that variables related to eyes off road time are of less importance than single glance duration for event detection and response performance, and likely cognitive distraction arising from secondary visual-manual tasks while driving.

Other studies have investigated the association of mean single glance duration with primary driving performance during secondary visual-manual tasks [24,27,28,29,30,31,32,33,34]. The PCA of the CAMP-DWM data and that of Young and Angell [2] are consistent in their results that mean single glance duration is associated with increased missed events and response times on the road, for some short duration secondary tasks that have relatively little driver workload.

4.5. SUGGESTIONS AND IMPLICATIONS

There are four practical suggestions and implications as to the applicability of these findings, and how these could be useful to the automotive safety engineer.
4.5.1. Use Event Variables as Well as Workload Variables in Driver Performance Safety Assessments

The first suggestion is that measures of event detection and response should be added to driver workload testing in driver performance safety testing of secondary tasks. Both types of measure are necessary to capture the major dimensions of driver performance. For example, using a single criterion comprising only variables loading on the driver workload dimension (such as “15 second rule” [35] for task times, or a 20 second guideline for “eyes-off-road time” [36]) will fail to find tasks that meet these driving safety criteria but still adversely affect driving safety because of poor event detection and response performance. This implication does not mean the 15-second rule or 20 second eyes-off-road time limit are invalid; they successfully provide an upper limit on the driver workload associated with tasks such as destination or address book alphanumeric entry. However, such driver workload criteria by themselves are incomplete (just as any single variable used in a two-dimensional framework would be incomplete). A single variable such as task time (or any other commonly used driver workload variable such as eyes-off-road time, number of glances, SDLP, etc.) does not capture the poor event detection performance of some low workload tasks as the current results show. A second variable or set of variables is needed to assess relatively low workload tasks as to their event detection and response performance.

Likewise, a task that meets event detection and response criteria alone will not ensure that the task meets driver workload criteria that could adversely affect driving safety (such as the destination entry task in the CAMP DWM study). Both dimensions must be considered conjointly, or secondary visual-manual tasks that can adversely affect relative crash risk will be missed. Considering both dimensions conjointly is necessary if task designs and guidelines are to improve event detection and response as well as driver workload, leading to optimal driver focus at all dimensions conjointly is necessary if task designs and driver performance safety testing of secondary tasks. Both such as the destination entry task in the CAMP DWM study were associated with P1. Therefore, neither P1 nor P2 can be a “visual distraction” dimension, because visual variables load on both dimensions. Likewise, P2 cannot be a purely “cognitive distraction” dimension, because it loads heavily with a visual variable as well as event detection and response variables. Therefore, this study suggests that driver distraction may be better segmented into “driver workload” distraction and “event-related” distraction. It is conjectured that cognitive distraction may be more closely associated with P2 than P1, because “cognitive capture” from some low-workload visual-manual tasks may give rise to long dwell times on a secondary device, reducing the event-related performance for these low workload tasks. Effects other than cognitive distraction may affect event detection and response, so the more general term “event-related” distraction is preferred than “cognitive distraction” to describe P2.

4.5.3. Do Not Combine Glance Duration and Number of Glances into Eyes-Off-Road Time

Eyes-off-road time is the product of task-related number of glances and mean single glance duration. However, the current analysis shows that the sole variable affecting eyes-off-road time is the number of glances. The current analysis further shows that the effect of long single glance duration on event detection and response is along a different dimension than is eyes-off-road time. This study shows that the effect of glance duration along the P2 dimension is “masked” when combined into a single eyes-off-road time variable, in agreement with Young and Angell [2]. Single glance duration and number of glances should be used as separate variables and not combined into eyes-off-road time. The Alliance Guideline 2.1 [36, p. 39] recognized the need for having two glance criteria: 1) “single glance durations generally should not exceed 2 seconds” and 2) “task completion should require no more than 20 seconds of total glance time to task display(s) and controls.” This study suggest that the current Alliance criterion of “total glance time” (i.e., eyes-off-road time) should be updated to a “number of glances” criterion in future guidelines.

4.5.4. Strengthen Single Glance Duration Guidelines

The fourth suggestion is that the criteria regarding single glance durations should be strengthened in future updates to automotive driver distraction and design guidelines. The current analysis shows that single glance duration is a key variable associated with event detection and response performance while driving. Long single glance durations are associated with increases in relative crash risk, more so than eyes-off-road time, as suggested by Horrey and Wickens [25], and substantiated in the 100-car data by Liang et al. [24], and Victor and Dozza [34] (see Section 4.4). Single glance duration effects may not even be linear. That is, a long single glance may have more effect on missed events and long reaction times than repeated short glances, even if the total eyes-off-road time is the same. That is another way of...
saying that the duration of single glances summates in a non-linear manner. This result was shown by Senders et al. [37] using an occlusion technique, where he stated that the uncertainty about the driving environment grows as a power function with an exponent of 1.5 as a function of not seeing the road (although more recent models question the 1.5 power exponent [24]). Although the technical details need further examination, these data suggest that single glance duration variables and criteria should be strengthened in future driver distraction guidelines.

4.6. LIMITATIONS

The tasks evaluated here were visual-manual only. The CAMP-DWM study by Angell et al. [6] found that auditory-vocal secondary tasks exhibited different driver performance profiles than visual-manual tasks. For example, many auditory-vocal tasks in the CAMP-DWM study had no eye glances off the road at all, causing all glance variables to task-related locations to be missing, and hence of no value for predicting driving performance. In addition, all of the auditory-vocal tasks had a pre-set task time (two minutes) in the CAMP-DWM study (except for BookOnTapeSummary). Hence, task time is not a useful variable for these tasks - task time cannot contribute to discriminating between tasks if it is experimentally held constant. Event detection and response metrics are among the few that can be validly compared between visual-manual and auditory-vocal tasks. Although PCA itself is completely general and can be applied to any data, a different set of variables than used here for analyzing visual-manual tasks will need to be developed for auditory-vocal tasks.

It remains to be investigated further the extent to which vehicle performance variables can be adjusted for task time by converting the raw variables into rates - rate of lane deviations, rate of speed deviations. Perhaps comparisons between visual-manual and auditory-vocal tasks can then be made more meaningful. While eye movements may not occur (or even be suppressed) in auditory-vocal secondary tasks while driving, it may still be the case that underlying attentional demands interfere with performance. This situation would be another case of “cognitive distraction,” and one that was not related to long single glance durations as with visual-manual tasks. Application of PCA methods to the driver performance variables associated with auditory-vocal tasks should be done in future research.

Another limitation is in the age range of the subjects. The subjects in the Young and Angell [2] study were all under the age of 65. There are order of magnitude differences in the mean data for subjects older than 65 in event detection performance during simulated driving compared to the mean data for those under 65 [38,39]. It is unknown if the driver performance dimensions described here will be the same in older subjects as for younger subjects. If they are not the same, then PCA methods applied to the separate age groups can identify the fundamental principal components for younger and older drivers, and more easily see the changes that occur. It is a lot easier to compare differences between tasks on two scores than the 15 variables in the Young and Angell [2] study, or in the 106 variables in the Angell et al. [6] CAMP-DWM study.

Another limitation is that only task means were evaluated. There may be important additional information contained in the “tails” of the distributions for single glance durations and reaction times (see Section 4.4). These additional variables can be analyzed in future studies now that PCA is known to work at least for the mean data in the major variables of the CAMP-DWM dataset as well as a previous dataset [1, 2].

5. SUMMARY

We attempted to define event detection and response (and cognitive distraction) on a more scientific basis, by an examination of the dimensions of driver performance in a key set of variables from the Crash Avoidance Metrics Partnership Driver Workload Metrics data set. The first and most prominent principal component explained about 53% of the total variance among the 10 variables, and related to conventional driver workload variables such as task time and eyes-off-road time. The second principal component explained about 25% of the remaining total variance among the ten variables, and related to event detection and response. Thus, ten variables could be reduced to two with a loss of only about one-fourth of the variance, the residual variance being mostly statistical noise. The two dimensions of driver performance found are orthogonal and independent.

Each task can then be scored on those two principal components. Instead of having to interpret the responses for ten driver performance variables for each secondary visual-manual task, the responses for only two variables need to be interpreted - the principal component scores. The scores on the principal components, when multiplied by the loadings on PC1 and PC2 respectively, can reproduce each task’s original data profile across all ten original variables to a high degree of accuracy. These two principal component scores can then be easily interpreted in a two-dimensional plot, instead of having to interpret ten variables that are highly intercorrelated (collinear), an almost impossible task.

It was confirmed that a task could be short with little workload (scoring low on dimension 1), and still have poor event detection and response (scoring high on dimension 2). Such task behavior was observed in the Young and Angell [2] study and noted by Angell et al. [2] in the CAMP-DWM study. It was referred to as a “paradoxical” finding in the CAMP-DWM study, but is easily explained if driver performance is considered in two dimensions rather than one.

Specifically, it is here suggested that the second principal component of driving performance may allow for a metric that can define event detection and response driving performance (and possibly also “cognitive distraction”), at least for visual-manual tasks. In addition, “mean single glance duration” was shown to be the glance metric that was of predictive value in estimating event detection and response...
performance (misses and response times to visual events during visual-manual tasks). On the other hand, the glance-related variables of eyes-off-road time and number of glances were associated with driver workload, but not with event detection and response. The separation of driver distraction into a visual distraction component and a cognitive distraction component is therefore misleading, because in fact glance variables play an independent and important role in both major dimensions of driver distraction. We therefore refer to dimension one as “workload distraction,” rather than “visual distraction,” and dimension two as “event-related distraction,” (which includes “cognitive distraction” as a special case). Vision variables are important to both dimensions, with different vision variables playing independent roles for each dimension.

The major results and implications were:

1. The principal component analysis (PCA) method applied to a subset of the mean visual-manual task data from the CAMP-DWM study track venue [6], confirmed the first two dimensions of driver performance as found by Young and Angell [2].

2. The similar results across the two studies show the robustness of these dimensions, because the two studies used entirely different events (one study using hood and side lights mounted on the vehicle administered repeatedly, and the other using brake light activation and deceleration of the lead vehicle administered once), as well as different visual-manual task sets.

3. The first component P1 was loaded by the five common driver workload variables of task time, glances, eyes-off-road time, standard deviation of lane position, and speed variations, and is identified as “driver workload.” P1 captured the majority of the variance in the CAMP-DWM data examined, as it did in Young and Angell [2].

4. P1 readily segments the tasks into low and high workload groups, but it failed to segment the relatively low workload tasks (those with relatively short task durations) into different groups. In particular, it failed to discriminate tasks with low vs. high miss rate and response times.

5. The second component P2 was loaded mainly by the event detection and response variables (miss rate and response time), and mean single glance duration. Mean single glance duration was the only glance variable among the three glance variables to load on P2. Tasks with high P2 scores exhibited longer glance durations and poorer event detection and response than those with low P2 scores.

6. This result supports the findings of Horrey and Wickens [25], and recent studies by Liang et al. [24] and Victor and Dozza [34], as well as earlier research regarding the key importance of long single glance durations as the predominant proximal cause of crashes, more so than measures that rely upon glance history variables.

7. The fact that eyes-off-road time and number of glances did not load onto P2 but only onto the orthogonal component P1, means that these glance variables have no predictive ability for event detection and response, under the conditions of the CAMP-DWM study and the set of variables examined here.

8. A new finding was that P2 readily segmented low workload tasks into separate groups with low, medium, or high miss rates and response times to visual events.

9. Because both P1 and P2 have vision-related variables associated with them, the two dimensions are perhaps better termed “workload distraction” and “event-related distraction” rather than “visual distraction” and “cognitive distraction.”

10. The long single glance durations observed in this study for tasks scoring high on P2 are conjectured to be the result of cognitive distraction causing “attention capture.” Certain short duration low workload tasks have yet undetermined design characteristics that give rise to long single glance durations and poor event detection and response.

11. PCA methods readily extend to other task types such as auditory-vocal, cell phone conversations, or “pure cognitive” tasks, or any other set of driver performance variables or venues. However, it is cautioned that the same set of variables used here for visual-manual tasks may not produce a meaningful result for these other task types. Hands-free auditory-vocal tasks do not have a definite control or display to look at, and therefore have missing values for the mean single glance duration for task-related glances, negating the usefulness of the long single glance duration variable for assessing event detection and response (and cognitive distraction) for such tasks.

6. CONCLUSION

The second dimension of driver performance for visual-manual tasks is associated with event detection and response (reaction times and miss rates to visual events) and long duration eye glances. It is orthogonal (uncorrelated) with the first dimension of driver workload, associated with the secondary task variables of task time, number of glances, subjective workload, lane variability, and other conventional driver workload metrics. Driver workload and “event detection and response” are therefore separate and unrelated dimensions of driver performance and distraction. These two categories provide a cleaner segmentation of driver distraction than “visual distraction” and “cognitive distraction.”

REFERENCES


DEFINITIONS/ABBREVIATIONS

CAMP-DWM
  Crash Avoidance Metrics Partnership Driver Workload Metrics

CHMSL
  Center High-Mounted Stop Lamp

EORT
  Eyes Off Road Time

FVTS
  Following-Vehicle Turn Signal

HVAC
  Heating, Ventilation, Air-Conditioning

LED
  Light-Emitting Diode

LVD
  Leading -Vehicle Deceleration

OED
  Object and Event Detection

P1
  Principal Component 1

P2
  Principal Component 2

PCA
  Principal Component Analysis

PDT
  Peripheral Detection Task

RT
  Response Time

SDLP
  Standard Deviation of Lane Position

VTTI
  Virginia Tech Transportation Institute

CONTACT INFORMATION

Richard A. Young, Research Professor
Department of Psychiatry and Behavioral Neurosciences
Wayne State University School of Medicine
9B-19 University Health Center
4201 St. Antoine, Detroit, MI 48201
ryoun@med.wayne.edu

ACKNOWLEDGMENTS

This research was supported in part by the Toyota Collaborative Safety Research Center in Ann Arbor, Michigan, USA. I thank Michael Posner, Bruce Weaver, Charlene Hallett, Linda Angell, and the anonymous reviewers for helpful comments on earlier drafts. A special thanks to Linda Angell for her insight and knowledge in the area of driving safety and the CAMP-DWM project, which has contributed substantially to the content of this paper.
APPENDIX

APPENDIX A. PRELIMINARY ANALYSIS OF TASKS WITH HIGH P2 SCORES

The three tasks with the highest scores on P2 are listed in Sections A.1, A.2, A.3 from highest to lowest. All these tasks had relatively low driver workload, poor event detection and long single glance durations. Preliminary task analyses from the CAMP-DWM report are given after the task descriptions in each section. These task analyses in the CAMP-DWM report were mainly along a single driver workload dimension. A rating using a verbal vs. spatial working memory emphasis was also given [6, p. 2-13]. The reader is referred to the original CAMP-DWM report [6, pp. 2-12 to 2-17] for more detail on these tasks and the 10 others analyzed here. The characteristics of these particular tasks (and others like them) that give rise to relatively poor event detection and response and long single glance durations, are unknown, and require further investigation.

A.1. HVAC TASK

The HVAC task (P2 score 2.4) made use of a conventional three-knob (fan, temperature, and airflow) climate control panel mounted on the top of the center stack area above the CD unit (Fig. A1).

The test participant was asked to adjust all three controls to desired levels using conversational language, e.g., “Your task is to adjust the heating, ventilation, and air conditioning unit so that the fan is high, at a moderately warm temperature, to warm both face and feet. Please begin now.” [7, p. B-3]. The task is visually (and probably cognitively) intensive because the sequence of mode changes the subject was asked to do was cleverly ordered differently than the controls on the HVAC interface (and the sequence in the task description), resulting in rapid searching before the three subtask commands were forgotten: 1) change mode to ‘face and feet’, 2) change fan to high, and 3) change temperature to moderately warm. The right knob had to be adjusted first, then the left knob, and then the middle knob (Fig. A1). This task also required short-term memory load, because the next knobs to adjust and their settings had to be remembered while executing the first and second subtasks. Angell et al. [6, p. 2-13] estimate by expert judgment that this task has “a spatial working memory emphasis” more than a “verbal working memory” emphasis in terms of the Multiple Resource Model of Wickens and Hollands [40]. The longer single eye glance durations than other tasks of similar task time might be because the ordering of the subtasks was different from the order of the knobs on the faceplate.

A.2. RADIOEASY TASK

The RadioEasy task (P2 score 2.1) was to tune the radio manually. The radio was already on, set to the appropriate band (FM) at a given preset station (100.1 FM). The test participant was asked to manually tune the radio to a specific frequency that was an approximately equal number of increments up or down from that setting (104.3 FM or 97.1 FM). They were to do this by means of the rotary knob provided for manual tuning on the device (see Fig. A2, [7, B.1.3]).
Figure A2. The radio was turned on and preset to 100.1 FM before the RadioEasy task [7, p. B-5].

None of the CAMP DWM radio tasks provided audio feedback, so drivers were more dependent on visual feedback for tuning. This task was designed to be similar to the radio tuning reference task that was initially specified in Principle 2.1B of the Alliance of Automotive Manufacturers Driver Focus Principles, Version 2.0 [36], but was slightly easier when implemented (it did not require turning on the radio, nor did it require selecting or changing bands, and required tuning slightly fewer increments to the target frequency) [7, B1.3]. Angelletal. [6, p. 2-13] estimate by expert judgment that this task has “a verbal working memory emphasis” more than a “spatial working memory” emphasis (the opposite of the HVAC task in A.1) in terms of the Multiple Resource Model of Wickens and Hollands [40]. It is not known why this particular task would have longer single eye glance durations than other tasks of similar task time, but perhaps long glances were required to discern the exact digits in the display, without the benefit of auditory feedback.

A.3. ROUTETRACE TASK

In the RouteTrace task (P2 score 2.0), the participant was asked to trace a path from a point of origin to a point of destination through a maze (Fig. A3). The mazes were 8 × 8 inches dimension and were created through a maze generator program. The task was intended to be analogous to the task of developing a route from a given location to a destination through surface streets and city blocks [6, p. B-15]. RouteTrace required near-continuous tracking of hand-with-pen tracing the maze. It is not known why this particular task had longer single eye glance durations than other tasks of similar task time, but it might possibly be related to the near-constant visual tracking required to trace the maze correctly.

Figure A3. Paper stimulus materials book ready for the maze task in vehicle [7, p. B-16].