

Event Detection: The Second Dimension of Driver Performance for Visual-Manual Tasks

Richard Young

ABSTRACT

A principal component analysis of the test track data from the Crash Avoidance Metrics Partnership Driver Workload Metrics project provided evidence for two major components in distraction during driving. The first was related primarily to driver workload, while the second was related to event detection and response. This result confirms previous test track findings. A new finding was that “mean single glance duration” (the average dwell time of the eye on the display or control needed to perform the task), loaded on the second dimension (associated with event detection and response), rather than the first (associated with driver workload). Hence, the duration of single glances to a secondary task is more important for event detection and response when driving than total eyes-off-road time or number of glances. These findings fit with the role of a single off-road glance immediately before a crash being predictive of crash probability. The usual division of “driver distraction” into “visual” and “cognitive” categories is misleading, and “driver distraction” is better subdivided into “workload distraction” and “event-related distraction.”

CITATION: Young, R., "Event Detection: The Second Dimension of Driver Performance for Visual-Manual Tasks," *SAE Int. J. Passeng. Cars - Electron. Electr. Syst.* 5(1):2012, doi:10.4271/2012-01-0964.

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 "glncedur" 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.

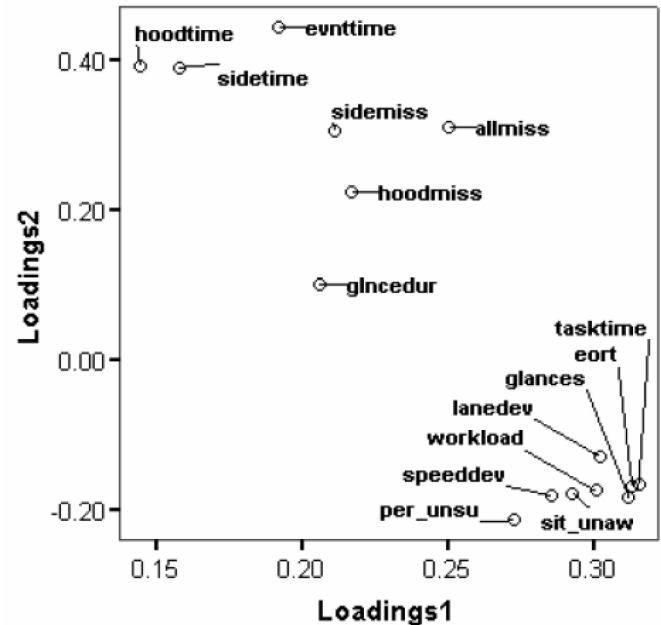


Figure 1. Two-dimensional plot of loadings for the first two principal components from the Young and Angell study [2, Fig. 2]. Used with permission.

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 [6] 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.

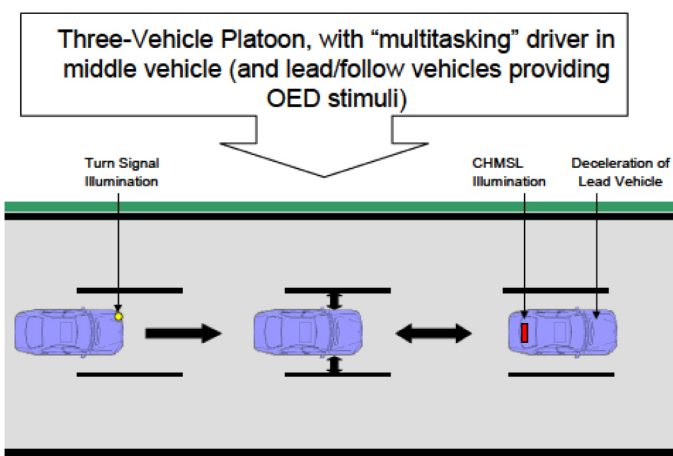


Figure 2. The event detection and response scenarios in the CAMP-DWM study [6, p. xxx].

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. [7, 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 “GlncDur” 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.

variable	CAMP-DWM [6] ^a	Appendix [7]	Young & Angell [2] ^a
1 TaskTime	Task duration	Q-34	Task duration
2 EORT	Mean duration of glances ^b	Q-45	Eyes-Off-Road Time ^a
3 Glances	Mean # glances ^b	Q-44	# Glances ^a
4 SDLP	Standard deviation of lane position ^c	Q-35	# Lane deviations ^j
5 SpeedDev	Speed difference ^d	Q-36	# Speed deviations ^k
6 GlncDur	Mean single glance duration ^b	Q-46	Mean single glance duration ^a
7 CHMSL%	% misses of CHMSL event ^e	Q-39	% Misses of hood event ^l
8 LVD%	% misses of LVD event ^f	----- ⁱ	% Misses of side event ^m
9 CHMSL_RT	Response time to CHMSL ^g	Q-38	RT to hood event ^l
10 LVD_RT	RT to LVD ^h	Q-37	RT to side event ^m

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.

TaskName	1 TaskTime	2 EORT	3 Glances	4 SDLP	5 SpeedDev	6 GlnceDur	7 CHMSL%	8 LVD%	9 CHMSL_RT	10 LVD_RT
ReadHard	30.9	15.3	13.7	0.64	5.73	1.04	18.8	26.9	2.26	5.76
Cassette	15.9	4.7	5.7	0.61	4.34	0.85	27.2	44.5	2.30	4.63
CDTrack7	18.4	9.8	9.1	0.63	4.79	1.03	15.8	46.3	2.19	5.49
ManualDial	24.4	11.0	11.3	0.68	5.67	0.97	30.2	24.5	2.30	5.94
Coins	17.6	6.7	7.3	0.65	4.52	0.90	26.4	51.2	2.39	5.94
ReadEasy	21.3	10.7	10.1	0.62	5.13	1.03	30.2	48.1	2.19	5.78
MapEasy	17.4	6.6	6.7	0.65	4.53	0.96	33.3	51.8	2.24	5.46
DestEntry	104.5	57.6	50.3	0.84	8.47	1.15	27.5	31.4	2.72	6.62
RadioHard	16.7	9.0	7.7	0.59	4.34	1.16	34.2	38.5	2.77	5.42
MapHard	25.2	12.7	10.7	0.74	5.63	1.10	32.2	46.5	2.61	6.83
RouteTrace	26.9	12.9	11.4	0.82	6.14	1.19	33.1	41.8	2.96	6.67
RadioEasy	11.4	5.7	4.6	0.52	3.73	1.28	33.3	72.7	2.38	5.93
HVAC	11.4	6.3	5.6	0.55	3.84	1.15	34.8	72.3	2.69	6.24

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 “GlnceDur” (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 $n = 13$.

	1 TaskTime	2 EORT	3 Glances	4 SDLP	5 SpeedDev	6 GlnceDur	7 CHMSL%	8 LVD%	9 CHMSL_RT	10 LVD_RT
1	1	.996	.998	.711	.914	.180	-.157	-.469	.302	.428
2	.996	1	.998	.684	.901	.248	-.147	-.434	.326	.457
3	.998	.998	1	.690	.906	.198	-.165	-.459	.298	.432
4	.711	.684	.690	1	.887	.046	-.049	-.574	.461	.599
5	.914	.901	.906	.887	1	.119	-.199	-.665	.314	.539
6	.180	.248	.198	.046	.119	1	.380	.313	.606	.573
7	-.157	-.147	-.165	-.049	-.199	.380	1	.401	.513	.273
8	-.469	-.434	-.459	-.574	-.665	.313	.401	1	.022	.010
9	.302	.326	.298	.461	.314	.606	.513	.022	1	.596
10	.428	.457	.432	.599	.539	.573	.273	.010	.596	1

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.

Variable	P1	P2
TaskTime	0.40888	-0.08138
EORT	0.40827	-0.05076
Glances	0.40705	-0.07730
SDLP	0.37305	-0.02450
SpeedDev	0.42068	0.11141
GlncDur	0.11795	0.49790
CHMSL%	-0.03773	0.49566
LVD%	-0.23503	0.37567
CHMSL_RT	0.20457	0.45521
LVD_RT	0.26809	0.36040

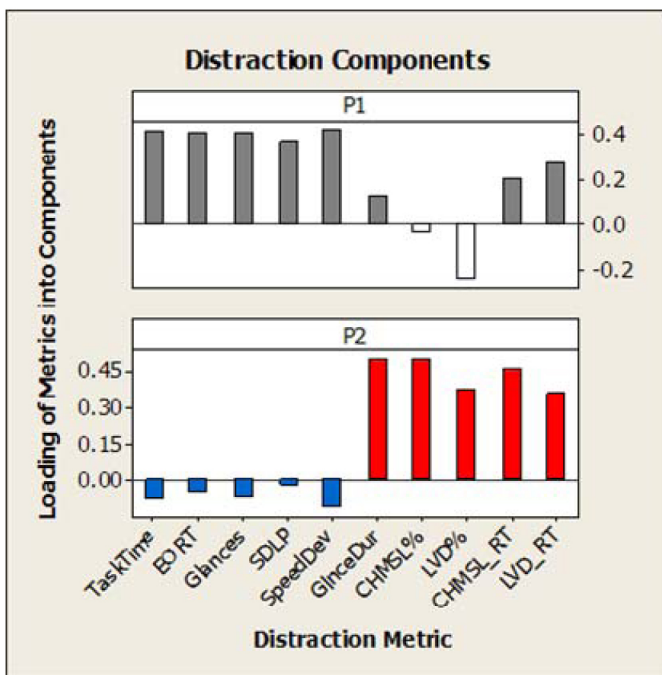


Figure 3. Loadings of the original variables on the first two components for the visual-manual tasks in the CAMP-DWM track study [6]. This is a plot of the variable loadings in Table 4).

Fig. 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 GlncDur is clearly grouped with the event detection and response variables in the upper left rather than the driver workload variables in the lower right.

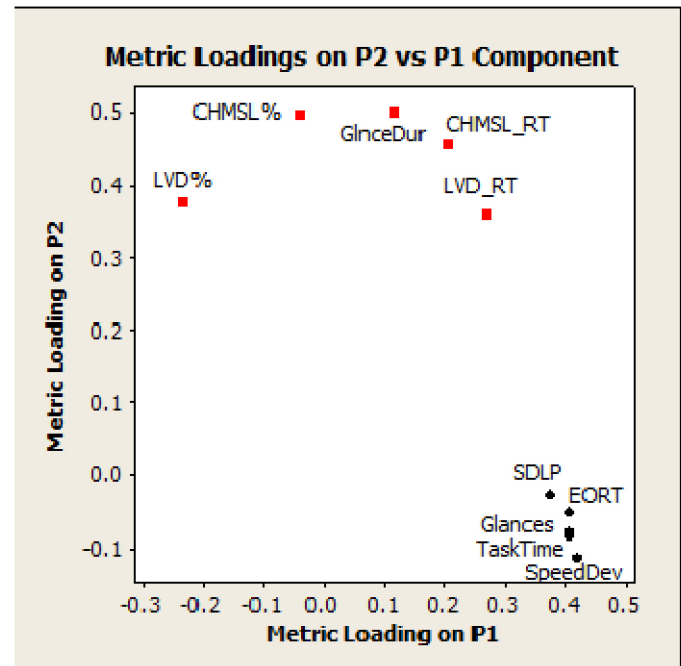


Figure 4. P2 vs. P1 variable loadings in Fig. 3 and Table 4 plotted against each other to show the associations between the variables in two dimensions.

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.

Task	S1	S2
ReadHard	0.4630	-1.9598
Cassette	-1.9544	-1.9002
CDTrack7	-0.9033	-1.8605
ManualDial	0.2934	-1.1174
Coins	-0.9893	-0.7033
ReadEasy	-0.7095	-0.4819
MapEasy	-1.3465	-0.4110
DestEntry	6.6562	-0.2750
RadioHard	-0.7470	1.0628
MapHard	0.9638	1.2055
RouteTrace	1.8757	1.9701
RadioEasy	-2.0335	2.0871
HVAC	-1.5687	2.3835