Driver Movement Patterns Indicate Distraction and Engagement

Robert G. Radwin, **John D. Lee**, and **Oguz Akkas**, University of Wisconsin–Madison

Objective: This research considers how driver movements in video clips of naturalistic driving are related to observer subjective ratings of distraction and engagement behaviors.

Background: Naturalistic driving video provides a unique window into driver behavior unmatched by crash data, roadside observations, or driving simulator experiments. However, manually coding many thousands of hours of video is impractical. An objective method is needed to identify driver behaviors suggestive of distracted or disengaged driving for automated computer vision analysis to access this rich source of data.

Method: Visual analog scales ranging from 0 to 10 were created, and observers rated their perception of driver distraction and engagement behaviors from selected naturalistic driving videos. Driver kinematics time series were extracted from frame-by-frame coding of driver motions, including head rotation, head flexion/extension, and hands on/off the steering wheel.

Results: The ratings were consistent among participants. A statistical model predicting average ratings from the kinematic features accounted for 54% of distraction rating variance and 50% of engagement rating variance.

Conclusion: Rated distraction behavior was positively related to the magnitude of head rotation and fraction of time the hands were off the wheel. Rated engagement behavior was positively related to the variation of head rotation and negatively related to the fraction of time the hands were off the wheel.

Application: If automated computer vision can code simple kinematic features, such as driver head and hand movements, then large-volume naturalistic driving videos could be automatically analyzed to identify instances when drivers were distracted or disengaged.

Keywords: naturalistic driving, video extraction, subjective rating scales, driver kinematics

Address correspondence to Robert G. Radwin, PhD, Department of Industrial and Systems Engineering, University of Wisconsin-Madison, 1550 Engineering Drive, Madison, WI 53706, USA; e-mail: radwin@discovery.wisc.edu.

HUMAN FACTORS
Vol. 59, No. 5, August 2017, pp. 844–860
DOI: 10.1177/0018720817696496

INTRODUCTION

Driver distraction represents a major safety problem in the United States; each day more than 15 deaths and over 1,200 injuries in crashes are reported involving a distracted driver (National Highway Traffic Safety Administration, 2010). The explosion of Web-based applications and connected vehicle information sources compete for drivers' attention and may make the issue even more critical in the coming years. In contrast to distraction, drivers' minds wander and they disengage from driving even in the absence of a competing activity. A case-control study based on data collected from crash survivors in a hospital emergency department found mind wandering in 17% of at-fault crashes, compared with only 9% in crashes where the driver was judged to have not been at fault (Galera et al., 2012). State and federal regulators lack sufficient understanding of the issue to develop regulations.

Naturalistic driving video provides a unique window into driver behavior unmatched by crash data, roadside observations, or driving simulator experiments. Data from such studies, such as the Strategic Highway Research Program 2 (SHRP2) Naturalistic Driving Study (NDS; Campbell, 2012), capture millions of hours of actual driving and hundreds of crashes but lack physiological data, reaction time probes, or subjective ratings that are commonly used to index internal states.

The objective of this study is to identify driver movements that are indicative of distraction and engagement behaviors. The motivation for this research is to develop a method to automatically review millions of hours of naturalistic driving video to identify likely occasions of driver distraction and lack of engagement. Linking distraction and engagement to overt behaviors is valuable because it would then be possible to use large-scale NDSs to assess the risk associated with these states. Because it is not possible to evaluate the driver's true mental state from a video depicting events that have already occurred, we investigate surrogate measures. We considered comparing overt behaviors, such as head and hand movements that can be objectively measured, against characteristics that observers of the video would subjectively identify as behaviors exhibiting distracted or disengaged driving. A computer algorithm could then be developed for tracking relevant driver body postures and motions so that distraction and engagement behaviors can be automatically classified accordingly. This study develops the concepts and methodology for such an approach and demonstrates it on a sample of naturalistic data.

Computer vision has previously been used for surveillance and characterizing human behaviors based on extracted movement properties (Gong et al., 2010; Hu, Tan, Wang, & Maybank, 2004). New computer algorithms are currently being developed to quantify high-level features pertinent to driver distraction and engagement in videos, like those in SHRP2 NDS (Smith et al., 2016), using a novel video analysis approach for tracking head position and estimating head pose and eye and mouth states. Our ultimate aim is to identify discrete representations of driver movement kinematics corresponding to distraction and disengagement behaviors for computer vision algorithms. In the current study, we develop the foundational research that relates manually extracted driver motions from naturalistic driving videos to subjective ratings of driver distraction and engagement behaviors.

Video feature extraction for recognizing affective states was previously used by Mota and Picard (2003) for associating postures and a child's interest level while performing a learning task on a computer. Jabon, Bailenson, Pontikakis, Takayama, and Nass (2011) associated extracted key facial features of drivers, using facial recognition software, with motor vehicle crashes. We developed empirical models that related hand kinematics properties of repetition frequency, hand speed, and duty cycle (ratio of exertion to rest time) to subjectively rated hand activity levels (Akkas et al., 2015; American Conference of Governmental Industrial Hygienists, 2009; Chen, Hu, Yen, & Radwin, 2013; Latko et al., 1997; Radwin et al., 2015). Successful implementation of computer vision depends on identifying relevant features to extract from the videos. Here we relate subjectively rated driver behaviors suggestive of various states of distraction and engagement to features of head and hand movements that a computer vision system might be able to extract from videos.

Definitions and Conceptual Model

We define *distraction* as the diversion of attention toward activities that compete with activities critical for safe driving (Lee, Regan, & Young, 2008). Examples of competing activities are eating, talking, adjusting music, smoking, texting or making a phone call, or operating the navigation system while driving. Distraction levels range from a driver who is not distracted at all and all effort is on driving to a driver who is fully distracted and no effort is on the road or on driving tasks while performing activities other than safe driving. Low distraction means no effort is on competing activities, whereas high distraction means all effort is on competing activities.

We define *engagement* as the level of effort devoted on all activities during driving (Yanko & Spalek, 2014), including competing and noncompeting activities. Engagement ranges from total neglect of attention to the road and driving tasks while the mind wanders to a fully engaged driver who is actively concentrating on driving and the roadway. Examples of various levels of concentration include sleeping, mind wandering, disinterest, paying attention to surroundings, or attentive driving. Multitasking, such as eating, texting, or conversation, can occur with various levels of engagement. Drivers can be highly engaged and intensely focused on the road and a secondary task, or they can devote relatively little attention to either the road or secondary task.

We conceptualize that levels of distraction and engagement are independent, are orthogonal, and occur on continuous scales. The level of distraction versus the level of engagement while driving is depicted in Figure 1. In this theoretical model, we define attentive driving as a combination of low levels of distraction and high levels of engagement, which occurs in the second quadrant. During attentive driving, many mental resources (Wickens, 2008) are deployed, and most are dedicated to driving. Conversely, inattentive driving is

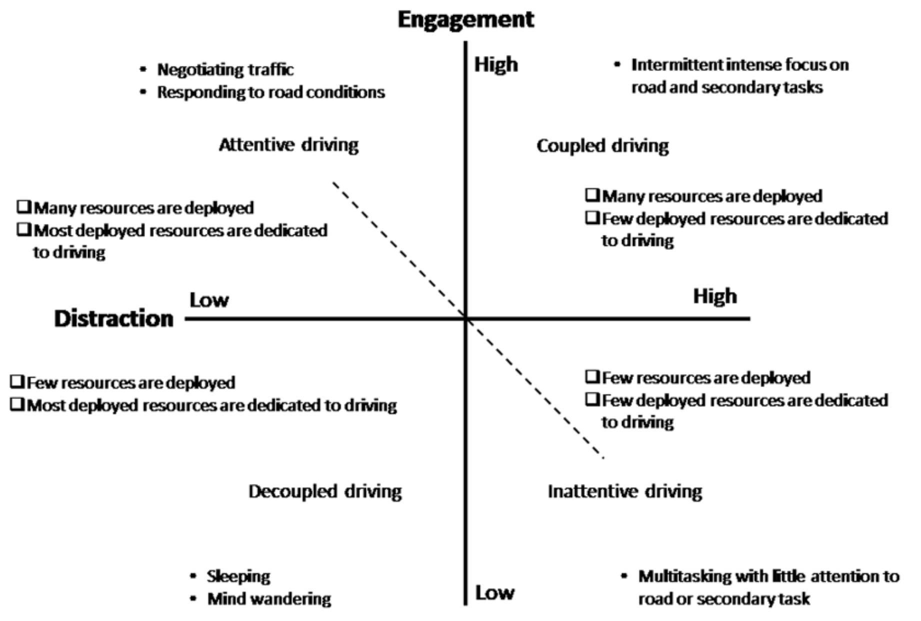

Figure 1. Driver distraction versus engagement.

a combination of high levels of distraction and low levels of engagement, which occurs in the fourth quadrant. When a driver is inattentive, few mental resources are deployed and few are dedicated to driving tasks; rather, most are focused on competing tasks, such as texting on a cell phone. Furthermore, we theorize that a decoupled state (third quadrant) is a combination of low levels of distraction and engagement, where only few resources are deployed although most are dedicated to driving, such as in mind wandering. The converse is that a coupled driver (first quadrant) has many resources deployed, but few are dedicated to driving; rather, most are dedicated to numerous competing tasks. We hypothesize that these states are a combination of levels of distraction and engagement, dependent on the quantity of mental resources used and the tasks to which they are committed. Thus, a mind-wandering or fatigued driver is disengaged but not distracted, whereas a multitasking or enraged driver is highly engaged while also distracted by competing tasks. We anticipate that a driver can move from one degree of distraction and engagement to another while driving.

Context has an important influence on the prevalence and consequence of mind wandering and distraction. Familiarity seems to precipitate disengagement and decoupled driving because well-learned patterns and habits make it possible for the driver to drive without full engagement, at least during routine situations (Yanko & Spalek, 2013). Likewise, benign driving situations encourage drivers to engage in secondary tasks. Any safety consequences of distraction and disengagement depend on how well drivers shift their attention to accommodate changes in the driving context. The current research focuses on simply quantifying observed driver state and not the more complex issue of its relationship to the driving context.

There is considerable evidence suggesting that driver distraction and engagement have characteristic movement behaviors. Lee, Oh, Heo, and Hahn (2008), using infrared sensors, found that head movement features can distinguish normal and drowsy drivers. While distracted, the proportion of time in which the driver's gaze is directed toward the road diminished, the proportion of head movements increased, and changes in gaze position upward increased (Metz & Krueger, 2010). Eye and head movements during driving are closely related (Fridman, Lee, Reimer, & Victor, 2016). Similarly, mouth movements can indicate distraction and fatigue (Wang, Guo, Tong, & Jin, 2004). Mou-

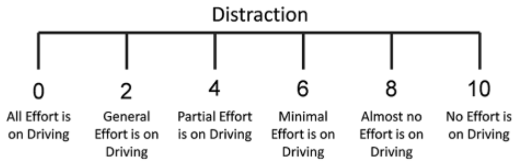

Figure 2. Distraction is the driver's level of effort on driving rated on a 0-to-10 visual analog scale.

rant and Rockwell (1970) early reported that drivers' search and scan patterns reflected differences in route familiarity and the type of driving conditions. He, Becic, Lee, and McCarley (2011) observed that during mind wandering, participants tended to scan the environment more narrowly. More generally, roadway fixation time was associated with driving performance (Chiang, Brooks, & Weir, 2004).

It is hypothesized that attentive driving (low-rated distraction and high-rated engagement behaviors) is associated with driver behaviors indicative of frequently scanning the road while the head elevation (neck flexion/extension) is oriented to the road (i.e., head rotation from left to right and less head elevation). Alternatively, inattentive driving (high-rated distraction and low-rated engagement behaviors) is associated with driver behaviors indicative of infrequent head scanning while gazing in directions other than the road (i.e., little head rotation from left to right and head elevation away from the front).

The purpose of this study is to consider how driver movement patterns are related to subjective ratings of driver distraction and engagement behaviors using feature extraction from video clips of naturalistic driving. Subjective rating scales were developed for observers to evaluate a driver's distraction and engagement levels when presented with short (15-s) NDS video clips. These scales were first demonstrated to participants who were asked to use them to rate a set of randomly presented video clips. Independently, the video clips were coded by an analyst frame by frame for specific driver motions, including head rotation, head flexion/extension, and hands on/off the steering wheel. Regression models were developed for predicting the participant ratings from the distraction and engagement rating scales based on kinematic variables for each clip to identify kinematic variables associated with ratings of distraction and engagement. These models will be used to input driver movements and output the ratings to identify portions of an NDS video where the driver is likely distracted or disengaged. The regression models were validated by having participants rate a different set of video clips that were similarly coded by an analyst for driver motions. The subjectively rated distraction and engagement values were then compared against the values predicted by the regression models.

METHOD

Rating Scales

Subjective rating scales (Annett, 2002) were created for evaluating driver distraction and engagement behaviors. Participants rated distraction behavior on a 0-to-10 visual analog scale, where 0 represents no distraction and "all effort is on driving" and 10 represents full distraction where "no effort is on driving." The distraction scale is illustrated in Figure 2. This scale rates the observer's appraisal of the driver's attention to driving while occupied with competing activities. We contend that driver distraction is related to the proportion of resources deployed that are dedicated to the driving task. When a driver's mental resources are diverted away from activities critical for safe driving by competing activities, we theorize that fewer mental resources are dedicated to driving and that this affects physical movements and postures assumed by the driver.

Engagement is the driver's level of concentration on all activities, where low engagement means no concentration and high engagement means full concentration. We rated engagement behavior on a 0-to-10 visual analog scale, where 0 represents the observer's appraisal that the driver's behavior indicates he or she is not engaged and "devoid of concentration" and 10 represents the observation that a driver behavior indicates he or she is fully engaged with "complete concentration." The engagement scale is illustrated in Figure 3. This scale rates the observer's evaluation of driver's level of concentration on all activities, based on the behaviors displayed.

Human participants were recruited to apply the rating scales for distraction or engagement

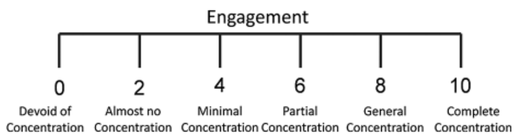

Figure 3. Engagement is the driver's level of concentration on all activities rated on a 0-to-10 visual analog scale.

behaviors after viewing video clips of naturalistic driving. The clips were taken from training data sets of the SHRP2 NDS (InSight SHRP2 NDS, 2016) videos. The SHRP2 NDS video data include detailed vehicle state data; the video record of the driver and surrounding road situation often provides a more revealing account of driver behavior (Campbell, 2012). The specific video clips shown were reviewed and selected by the experimenters for depicting various driving activities, such as attentive driving, talking, using a telephone, or yawning, while attempting to select a range of distraction and engagement levels.

This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at the University of Wisconsin–Madison. Informed consent was obtained from each participant.

Benchmark Practice Videos

A subset of six SHRP2 NDS video clips were reserved for practice using the rating scales. These clips (approximately 15 s each) were all rated for distraction and engagement behaviors by a consensus panel of eight researchers to provide participants with examples of the ratings and for learning to use the scales. The researchers together viewed each video clip and then independently rated it according to either the distraction or the engagement behavior scales shown in Figures 1 and 2. Video clips were displayed by playing three simultaneous views of the driver's face, hands, and front view (Figure 4). After individually rating a clip, each researcher recorded and then announced his or her rating. If there was consensus, defined as ±1 unit ratings, then the average rating was applied. Any nonconsensus ratings were discussed, and each explained why he or she gave the score he or she did. Typically, this discussion occurred if something in the video clip was overlooked. The video was then rerated until there was consensus. The benchmark video clips included distraction ratings from 0.9 to 5.4 and engagement ratings from 4.0 to 8.9, which were rounded to the nearest integer.

Naturalistic Driving Videos Ratings

Twenty participants were recruited from the University of Wisconsin–Madison campus under informed consent. Inclusion criteria were having a current active driver's license from the United States, at least 3 years' experience driving, having driven in the past 6 months, never having had a driver's license suspended, student status or having recently graduated, and having taken and passed the Collaborative Institutional Training Initiative (CITI Program at the University of Miami, 2013) social and behavior science human subjects training, as some video clips potentially depicted actual human participant drivers.

All participants received practice on the distraction and engagement behavior rating scales. The scales in Figures 1 and 2 were introduced, and participants were not instructed about which features to observe other than to rate the level of distraction and engagement attributed to the driver they observed in the video clips. Participants were asked to view each clip carefully (Figure 4), which was played three times, and to rate it according to the respective scale based on their evaluation of their observations of the driver's behaviors. As part of practicing use of the rating scales, a benchmark video was played, the participants rated it, and they were then informed of the consensus panel ratings. The participants were then shown an additional clip and were asked to rate them using either the distraction or engagement behavior rating scales. Their rating was compared against the consensus panel rating. Practice ended after the participants demonstrated that they rated the benchmark clips in accordance with the panel within ±1 unit. This typically occurred after viewing two or three clips for each rating scale. All participants achieved this level of performance after practicing on less than the full set of six video clips.

A set of 36 selected SHRP2 NDS video clips (referred as the naturalistic driving videos), each

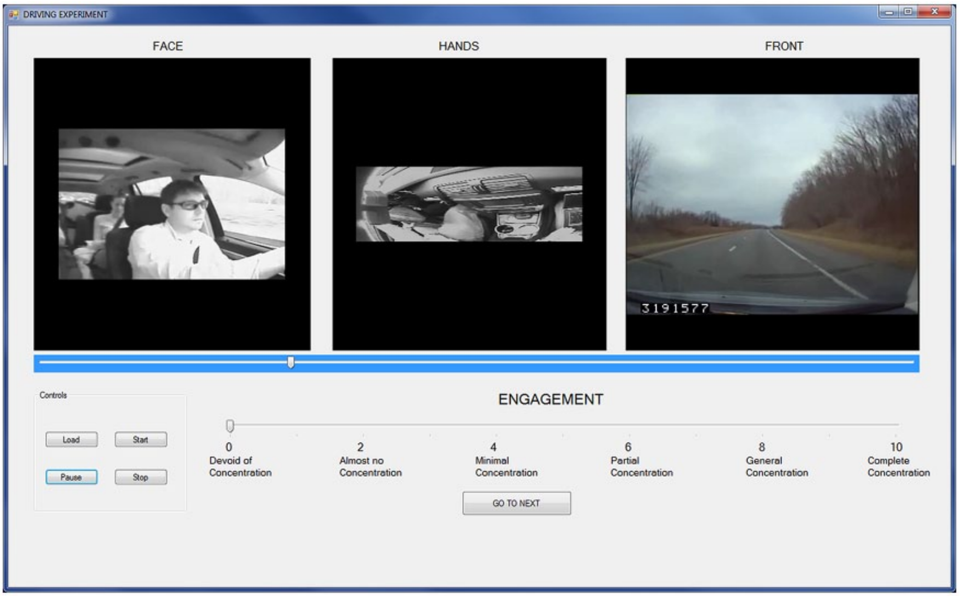

Figure 4. Example of screen shown to participants rating distraction and engagement, depicting the driver's face, hands, and front view from the vehicle.

about 15 s in length, was then shown to the participants during each session in random order. Participants were instructed to view each clip carefully in its entirety, which was played three times, and to rate it according to the assigned scale for that session. Participants were not informed about what features to consider in making their ratings, other than to carefully view the video. Participants rated all the naturalistic driving video clips once each session, for a total of four sessions conducted on different days. Two sessions were used to rate distraction behavior, and two sessions were used to rate engagement behavior, for a total of 1,440 ratings (20 participants × 36 video clips × 2 sessions) on each rating scale.

The average and standard deviation ratings of all of 20 participants were calculated for each video clip. If an individual rating of a clip was less than the 0.10th percentile or exceeded the 99.9th percentile of a normal distribution based on the mean and standard deviation of that clip, it was considered an outlier and was removed. A total of five ratings (0.35%) for distraction and seven ratings (0.49%) for engagement were removed (12 ratings in total); not more than one outlier rating occurred for any video clip.

Test–retest reliability was calculated by regressing all of the initial ratings for each of the video clips (20 participants × 36 video clips = 720 data points) against the corresponding ratings from a repeated data collection session. Interrater reliability was measured using Fleiss's Kappa and single-score intraclass correlation.

Extracted Movement Features

The 36 naturalistic driving face and hands view (Figure 4) video clips were also manually classified frame by frame by an analyst who manually coded driver position states, including head rotation, head flexion/extension, and hands on/off the steering wheel. Five symmetric head rotation and three symmetric head flexion/extension states were estimated (McAtamney & Corlett, 1993). Head rotation states were coded as 0° forward (−22.5° to +22.5°), −45° left (−22.5° to −67.5°), −90° far left (−67.5° to −112.5°), 45° right (+22.5° to +67.5°), and 90° far right (+67.5° to +112.5°). Head flexion/extension states were coded as 0° neutral (−7.5° to +7.5°), 15° extend up (+7.5° to +22.5°), and −15° flex down (−7.5° to −22.5°). Left-hand and right-hand states were each coded on or off.

TABLE 1: Extracted Feature Quantification Variables

Kinematic Variable	Description	Calculation
Head motion zero crossing	The frequency of head rotation (yaw) or flexion/extension (pitch) per second	Count each time head moves away from forward (0°), divided by the duration of the video clip
Head motion variance	Variability of head rotation (yaw) or flexion/extension (pitch) angle	Variance of the head angle per frame i.e., $\sum(\text{angle} - \text{mean angle})^2/n$
Area-under-head motion	Magnitude of head rotation (yaw) or flexion/extension (pitch)	Trapezoidal rule to estimate area under angles over the duration of the video clip
Percentage time not facing forward	Percentage of the time the head is not facing forward	Total time the head is not forward, divided by total duration of the video clip
Hands off the steering wheel	Percentage of the time the left or right hand is off the wheel	Total time the hand is off the wheel, divided by total duration of the video clip

The coding involved displaying the driver view of video clips one frame at a time and classifying the head and hand states for every frame using Multimedia Video Task Analysis™ software. The result was a time series of the driver position states, sampled every 33 ms, distributed over the duration of each video clip. The coding accuracy and reliability were checked by having another experimenter recode four randomly selected video clips. The average difference was 7.3° for head rotation angle and 2.3° for head flexion/extension angle. The average difference was −0.1% for right hand off the wheel and 7% for left hand off the wheel.

The coded frame-by-frame kinematic data were analyzed as a one-frame (1/33 ms = 30.3 Hz) sample rate time series. Features based on each of the states were calculated from the respective frame-by-frame time series for each of the 36 naturalistic video clips. Frequency of head motions was quantified by counting the rate of zero crossings during the time of the video clip. The magnitude of head motions was the absolute area under the signal, and variation of head motions was defined as the variance. The duration that the head was pointing away from forward was calculated as the percentage of time the head was not in the forward direction for the duration of the clip. The coded driver posture states were quantified using the features described in Table 1 and then compared against the distraction and engagement scales.

Modeling Subjective Ratings From Extracted Features

Participant ratings from the distraction and engagement scales were regressed against the data reduced kinematic variables described in Table 1 for every video clip in order to identify kinematic variables associated with ratings of distraction and engagement behaviors. We applied Akaike information criterion (AIC) as a means for regression variable selection. AIC is defined as $2k - 2\ln(L)$, where k is the number of parameter and $\ln(L)$ is the log likelihood of the model. In each step we chose a parameter (variable) to exclude for the model and compared AICs of the new model versus the previous model. We then selected the set of parameters that produced the smallest AIC. After the minimum AIC was obtained, the reduced models were subsequently reiterated using p value reduction ($p < .10$) to avoid spurious correlations and overfitting the data.

Model Validation

The regression models predicting rated distraction and engagement behavior from the extracted driver motion variables were validated by rating an independent set of naturalistic driving video clips (referred to as the validation driving videos) for distraction and engagement behavior (Figures 2 and 3), coding them for kinematic head and hand movement features, as described earlier, and entering them into the previously developed regression model. Ten participants were recruited to observe and rate 168 SHRP2 NDS (Campbell, 2012) validation driving video clips for distraction and engagement behaviors using the same scales and procedures as previously described. These clips contained similar views and were of quality as the previous set of naturalistic driving video clips. All the participants practiced on the scales using the same benchmark set of videos as the previous experiment. The validation videos were rated in two sessions, one for engagement and one for distraction, selected in a random order. Each session took 2 to 3 hr to complete.

It was not practical to manually code all the driver movement states for all the 168 rated validation video clips frame by frame, so a sampling process was employed. The rated video clips were first categorized based on their average distraction and engagement behavior ratings, rounded to the nearest integer, from all 10 participants. Because the distribution of mean ratings among the 168 videos was not uniform, we randomly selected clips from each rating category. Selected videos containing nondriving activities, such as stopping at a light, were excluded, resulting in 15 video clips for distraction (range 0–6.3) and 19 clips for engagement (range 5.3–8.6). Individual ratings outside of the lower 10th percentile or above the 90th percentile of the average were excluded.

The selected video clips were then manually coded for driver posture states as previously done for head rotation, head flexion/extension, and hands on/off the steering wheel. The frame-by-frame coded driver posture states were quantified using the features described in Table 1 and were entered into the regression model. The predicted ratings were then compared against the median participant ratings for each clip.

RESULTS

Rating Scales for the Naturalistic Driving Videos

Linear regression for average test–retest ratings among the 20 participants had a 0.84 coefficient and 0.42 intercept (R^2 = .92) for rated distraction behavior and a 0.92 coefficient and 0.83 intercept (R^2 = .91) for rated engagement behavior, which indicated good consistency in aggregate. Although test–retest reliability on average was good, the raters demonstrated individual bias. Linear regression for test–retest among all individual participant ratings had a 0.74 coefficient and 0.80 intercept (R^2 = .58) for rated distraction behavior and a 0.59 coefficient and 2.9 intercept (R^2 = .34) for rated engagement behavior. Fleiss's Kappa was 0.052 (p < .001) for distraction and 0.039 (p < .001) for engagement. The single-score intraclass correlation was 0.33, F(35, 760) = 24.3, p < .001, for rated distraction and 0.23, F(35, 761) = 15.7, p < .001, for rated engagement. Consequently, the participants strongly agreed on average, but less with each other.

The distribution of distraction and engagement behavior ratings for the naturalistic driving videos, averaged over all participants and repeated data collection efforts, is plotted in Figure 5. The correlation between the average distraction and engagement ratings was −0.47, t(34) = 3.4, p < .001, indicating a moderately strong association between distraction and engagement for the videos selected for the current study. Average distraction ratings ranged from 1.2 to 6.2, and average engagement ratings ranged from 4.7 to 8.6.

Subjective Rating Models From Naturalistic Driving Video Kinematics

Coded features for head rotation, head flexion, and hands on/off the wheel signals were quantified based on the calculated data described in Table 1. The relationships between the average kinematic variables and average distraction and engagement ratings are plotted for head rotation, head flexion/extension, and percentage of time the hands were off the wheel in Figure 6, 7, and 8, respectively.

Using stepwise regression, the percentage of time that the right hand and left hand are off the steering wheel were significant predictors of both rated distraction and rated engagement (Table 2). The sign for the hands-off coefficients was positive for the distraction model, increasing distraction when the hands were off the steering wheel more, and negative for the

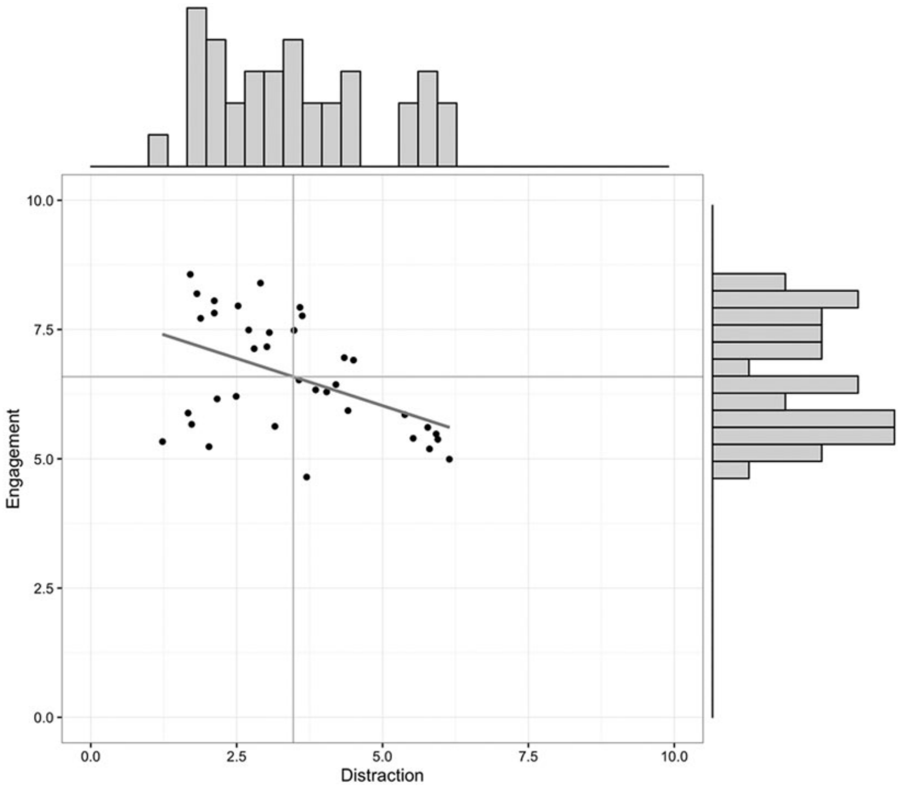

Figure 5. Average distraction ratings plotted against engagement ratings for each video clip. The relative distributions of distraction and engagement are shown above their respective axes.

engagement model, increasing engagement when the hands were more on the steering wheel. In addition, head rotation was a significant predictor of distraction, whereby greater head movements increased distraction (i.e., area-under-head rotation signal), and the variance of head rotation was a significant predictor of engagement, whereby more head scanning increased engagement (Table 2). A summary of the predictive models for distraction (R^2 = .54, *MSE* = 1.05), $F(3, 32)$ = 12.4, $p < .001$, and engagement (R^2 = .50, *MSE* = 0.85), $F(3, 32) = 10.51$, $p < .001$, are provided in Table 2.

Model Validation

After entering the extracted discrete kinematics data from the independent validation video clips into the regression models described in Table 2, predicted distraction and engagement ratings were plotted against the average participant ratings of distraction and engagement for those clips in Figure 9. The slope and intercept for distraction were 0.35 and 2.24, respectively, $F(1, 18) = 10.45$, $p < .01$, and the slope and intercept for engagement were 0.34 and 4.30, respectively, $F(1, 17) = 2.86$, $p < .11$, when a linear regression line was fitted.

DISCUSSION

Aggregate participant-rated distraction and engagement behavior on average was consistent when repeatedly viewing the video clips (R^2 = .92 for rated distraction and R^2 = .91 for rated engagement). The distraction and engagement behavior ratings were somewhat less consistent relative to other participants (R^2 = .58 for rated distraction and R^2 = .34 for rated engagement). Kinematic measures predicted subjective ratings of distraction and engagement behavior (R^2 = .54 for distraction and R^2 = .50 for engagement).

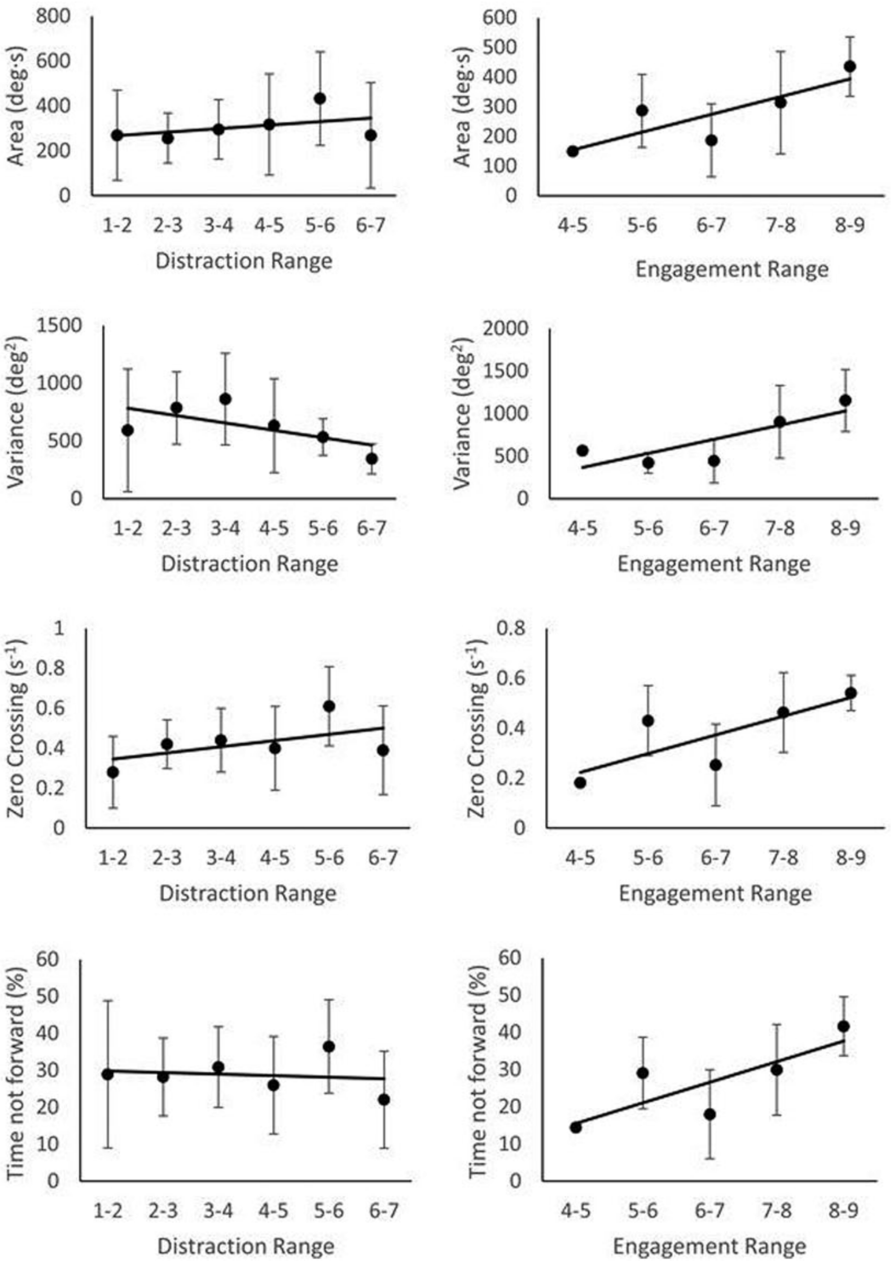

Figure 6. Area-under-the-head rotation signal, head rotation variance, head rotation zero-crossing frequency, and percentage time not facing forward are plotted against average distraction and engagement within a range of ±1 rating unit. Error bars are ±1 standard error.

Overall, the results suggest that behaviors of the driver can be rated and predicted.

For distraction behavior, a subjective rating of a driver's dedication to competing activities, positive relationships were observed between distraction behavior ratings and kinematic measures for head rotation area (Figure 6) and proportion of right hand and left hand off the steering wheel (Figure 8), indicating that the more the head rotated or the more that

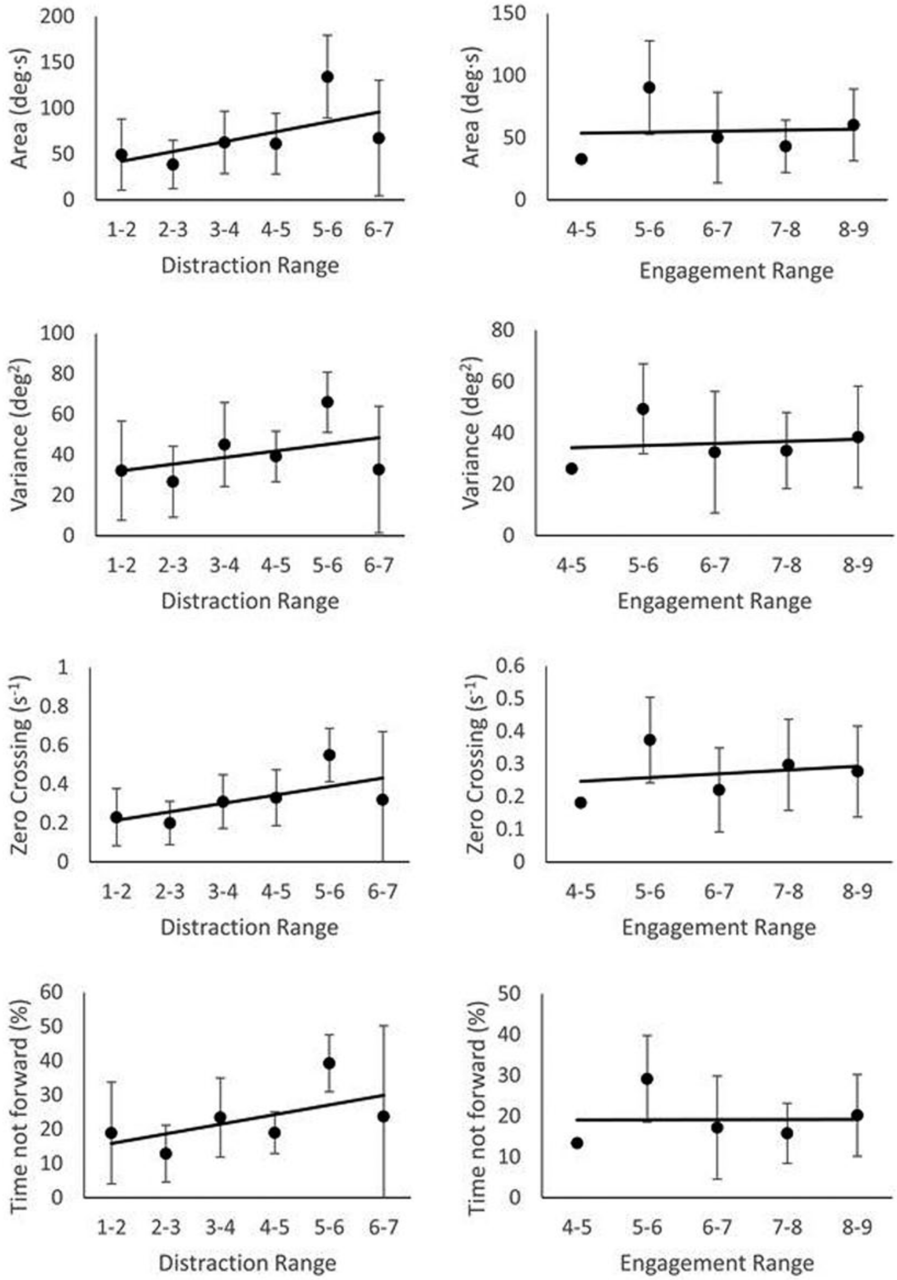

Figure 7. Area-under-the-head flexion/extension signal, head flexion/extension variance, head flexion/extension zero-crossing frequency, and percentage time head flexion/extension is not facing forward are plotted against average distraction and engagement within a range of ±1 rating unit. Error bars are ±1 standard error.

the hands were off the wheel, the higher the subjective ratings of distraction given. Positive relationships were observed between distraction behavior ratings and percentage of time the driver's head faces forward (Figure 6), meaning the more that the head faces forward, the lower the subjective ratings of distraction behavior given.

TABLE 2: Stepwise Regression Model Variables

Kinematic Variable	Coefficient	*SE*	*t* Statistic	*p* Value
Distraction				
Intercept	1.8209	0.3753	4.852	.001
Area-under-head rotation signal	0.0014	0.0008	1.769	.086
Right hand off steering wheel	0.0159	0.0043	3.703	.001
Left hand off steering wheel	0.0287	0.0063	4.572	.001
Engagement				
Intercept	6.5841	0.3275	20.103	.001
Head rotation variance	0.0010	0.0003	3.591	.001
Right hand off steering wheel	−0.0098	0.0036	−2.733	.010
Left hand off steering wheel	−0.0115	0.0051	−2.257	.030

We also found different kinematic measures associated with subjective ratings of engagement behavior, the subjective rating of a driver's concentration. Positive relationships were observed between engagement behavior and kinematic measures of head rotation variance (Figure 6), indicating that the greater the head rotation variation, the higher the subjective ratings of engagement behavior given. Negative relationships were observed between engagement behavior and kinematic measures for the proportion of right hand and left hand off the wheel (Figure 8), indicating that the more that the hands were off the wheel, the lower the subjective ratings of engagement given.

Although the regression models aptly predicted ratings of distraction and engagement behavior, the model R^2 values of .54 for distraction and .50 for engagement indicate that there was a considerable unaccounted variance. This finding may be partially attributed to the low-precision head motion variables used in the model. We anticipate higher precision when using computer vision video extracted variables, which should have a better signal-to-noise ratio and therefore introduce less random variation. Another source of the unaccounted variance may be the imperfect agreement of the subjective ratings. The regression models predict aggregate ratings, which are measures of the underlying driver state, and so make perfect prediction impossible.

Using observable measures to infer underlying mental states represents a fundamental challenge, even with physiological measures of brain activity and verbal reports (Diener, 2010; Mather, Cacioppo, & Kanwisher, 2013; Nisbett & Wilson, 1977). Such data are impossible to collect in a naturalistic driving context, and so we are left to use much more distal measures to estimate safety-critical driver states. Attentive driving is a combination of low levels of distraction (i.e., few competing activities) and high levels of engagement (i.e., focused attention to driving) and is depicted in Figure 1 in the second quadrant. During attentive driving, many mental resources are deployed and most of these deployed resources are dedicated to driving. This state may be observed by higher levels of head motion while frequently scanning the road as seen by head rotation variance. We therefore extend our conceptualized model in Figure 1 as body movement behaviors described in Figure 10. Any association between kinematic features and driver state must be viewed as tentative and imprecise; however, it might also offer a powerful means of identifying situations that are associated with increased crash risk.

Correspondingly, inattentive driving is a combination of high levels of distraction and low levels of engagement, as depicted in the fourth quadrant of Figure 1, where few mental resources are deployed, and few of the deployed resources are dedicated to driving activities. This state may be observed by infrequent head scanning as seen by less head rotation as described in Figure 10. The distribution of distraction and engagement levels in the video clips

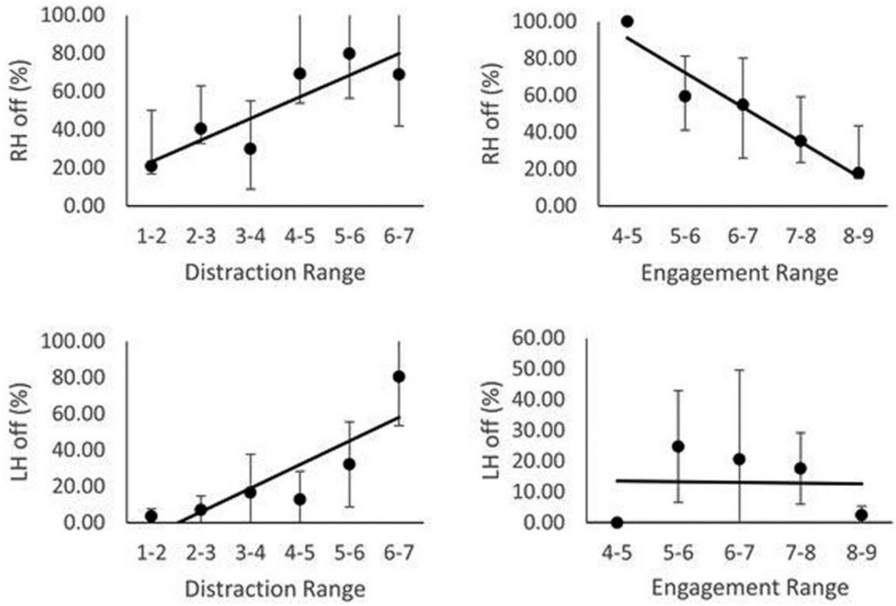

Figure 8. Percentage of time the left hand and percentage of time the right hand is off the steering wheel, plotted against average distraction and engagement within a range of ±1 rating unit. Error bars are ±1 standard error.

selected for the current study does not allow us to test our hypothesis for the remaining combinations (Figure 5).

Authors of future research should evaluate driver states using the scales developed in this paper possibly through driving simulator studies, whereby drivers self-evaluate distraction and engagement in comparison with independent observer ratings. In addition, objective physiological measures, such as electroencephalogram and eye tracking, might be employed and correlated against subjective distraction and engagement ratings.

Although a moderately strong correlation was observed between distraction and engagement behavior ratings (−.47), it is important to recognize that the video clips originated from a selected set of videos obtained from convenience and therefore did not represent a particular distribution of events or random sample of driving activities. Consequently, the correlation merely represents the relationship between distraction and engagement for the videos that were selected and does not indicate how they are related under driving conditions for any population of drivers. This relationship must be explored for a representative set of naturalistic driving videos. Similarly, the video clips used for validation (Figure 9) were limited in range, particularly for engagement. The predicted engagement pattern suggests a strong relationship, but this relationship is likely a by-product of the range restriction.

The distribution of ratings for the scales was dependent on the events that occurred in this limited set of naturalistic driving videos made available for this study. This study was a proof-of-concept investigation and does not represent the full spectrum of driving activities. The objective of this research was to demonstrate that observer ratings may be predicted with limited measures of driver positions and movement kinematics. We conclude that such a relationship can be observed.

Our framework describes two dimensions. One is the direction of attention toward or away from the road. This dimension is typically defined as distraction. The second is the amount of attention or degree of attention. The first conflates high and low effort, and the second conflates attention to the road or attention to the secondary task; however, in combination, they

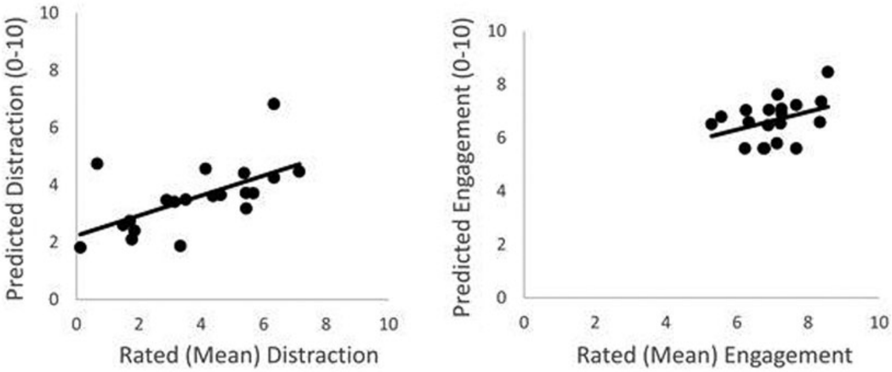

Figure 9. Predicted distraction (left) and engagement (right) plotted against mean ratings.

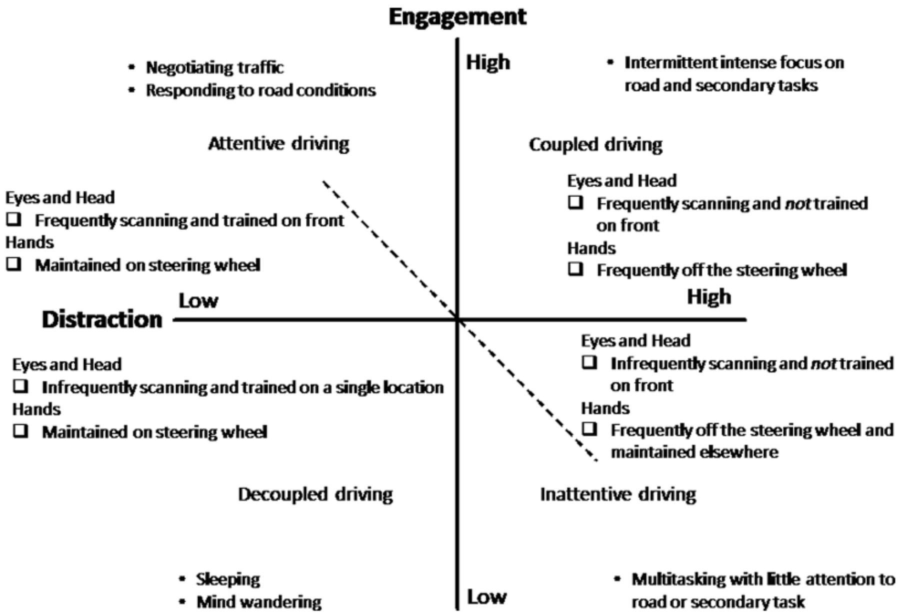

Figure 10. Model depicting relationship between driver behavior and coded kinematic features.

more fully map the space. Whether these distinctions can be easily measured is an empirical question and one that this study provides only an initial start at answering.

This paper shows weak support for the theoretical distinctions we propose. One explanation for this result is that the direction of attention is easier to judge from the video recordings than is the intensity of attention. We view this outcome as an important contribution because it shows possible limits on the method that should be considered in its future application. Certainly, the theoretical framework needs a series of studies to establish a more definitive basis.

Future research will extract additional driver movement features and postures, such as kinematic measures of head movement and movements of the hands. We are developing computer vision algorithms for automatically extracting driver movement states (Smith et al., 2016). Compared with the manually coded driver states, whereby driver posture states were limited to just a few visually estimated zones, we anticipate that automatic video extracted data will have better precision and therefore may provide improved models representing a greater proportion of the variance for the distraction-and-engagement-prediction model. Furthermore, applying the approach developed in this research for rating distraction and engagement might offer better reliability and more consistency than human raters.

CONCLUSIONS

This study shows distraction and engagement behavior can be rated in a consistent fashion and that these ratings can be predicted with kinematic data. We theorized that distraction and engagement behaviors can be independently rated, but we have not established conclusively if they are independent constructs that describe deployment of cognitive resource deployment during driving. The correlation between distraction and engagement (−.47) should be studied in the future using appropriately sampled naturalistic driving videos.

This study shows a promising approach to identifying potentially hazardous driver states from relatively impoverished naturalistic driving data in the absence of more proximal measures of cognitive processes, such as physiological measures and eye movements. We have shown that specific objective kinematic features were associated with subjective ratings of distraction and engagement. Although these relationships account for physical factors that divert attention from safe driving, such as redirecting vision away from the road, or that diminish control of the vehicle, such as removing the hands from the steering wheel, they are undoubtedly not fully indicative of the true cognitive states involved in driver distraction and engagement. We are simply, so to speak, reading the driver's body language. At a more pragmatic level, predictions of distraction and engagement might be used to estimate crash risk observed in naturalistic data. Such estimates of crash risk depend on how well drivers shift their attention to accommodate changes in the driving context. Using kinematic features to estimate driver state is a first step in extracting information from naturalistic driving data to understand the underlying cognitive processes that affect driving safety.

Additional physical measures, such as eyelid closing or mouth movements, might offer supplementary measures to help better associate kinematics with subjective ratings. These measures should at best help us get closer to revealing the specific movements and gestures that participants take into account when subjectively rating driver distraction or engagement but not decisively account for the actual cognitive resource deployment of individuals while driving. Future research might, through simulator studies, permit comparisons of distraction and engagement ratings from external participants against internal ratings made by the drivers in order to test if external movements observed are related to internal cognitive states.

ACKNOWLEDGMENT

This material is based upon work supported by the Federal Highway Administration under Contract No. DTFH6114C00011 as part of the Exploratory Advanced Research Program.

KEY POINTS

- The distraction scale subjectively rates a driver's performance in effort dedicated to driving when occupied with competing activities, and the engagement scale subjectively rates a driver's performance in concentration, and these ratings were consistent among viewers of naturalistic driving video clips.
- Positive relationships were observed between ratings of distraction behavior and kinematic measures for the area-under-head rotation signal and percentage of time that the hands were off the wheel, indicating that the more the head rotated or the more the hands were off the wheel, the higher the subjective ratings of distraction behavior given.
- Positive relationships were observed between engagement behavior and kinematic measures of head rotation variance, indicating that the greater the head rotation variation, the higher the subjective ratings of engagement behavior given.
- Negative relationships were observed between engagement behavior and kinematic measures for the percentage of time hands were off the wheel,

indicating that the more the hands were off the wheel, the lower the subjective ratings of engagement given.

REFERENCES

Akkas, O., Azari, D. P., Chen, C-H., Hu, Y. H., Ulin, S. S., Armstrong, T. J., Rempel, D., & Radwin, R. G. (2015). A hand speed and duty cycle equation for estimating the ACGIH hand activity level rating. *Ergonomics*, *58*, 184–194.

American Conference of Governmental Industrial Hygienists. (2009). Hand activity level TLVs® and BEIs®. In *Documentation of the threshold limit values for chemical substances and physical agents and biological exposure indices* (pp. 196–198). Cincinnati, OH: Author.

Annett, J. (2002). Subjective rating scales: Science or art? *Ergonomics*, *45*, 966–987.

Campbell, K. L. (2012). The SHRP 2 Naturalistic Driving Study: Addressing driver performance and behavior in traffic safety. *TR News*, *282*, 30–35.

Chen, C-H., Hu, Y. H., Yen, T. Y., & Radwin, R. G. (2013). Automated video exposure assessment of repetitive hand motion. *Human Factors*, *55*, 298–308.

Chiang, D. P., Brooks, A. M., & Weir, D. H. (2004). On the highway measures of driver glance behavior with an example automobile navigation system. *Applied Ergonomics*, *35*, 215–223.

CITI Program at the University of Miami. (2013). *Copyright*. Retrieved from http://www.citiprogram.org/index.cfm?pageID=260

Diener, E. (2010). Neuroimaging: Voodoo, new phrenology, or scientific breakthrough? Introduction to special section on fMRI. *Perspectives on Psychological Science*, *5*(6), 714–715.

Fridman, L., Lee, J., Reimer, B., & Victor, T. (2016). "Owl" and "lizard": Patterns of head pose and eye pose in driver gaze classification. *IET Computer Vision*, *10*, 308–313.

Galera, C., Orriols, L., M'Bailara, K., Laborey, M., Contrand, B., Ribereau-Gayon, R., . . . Lagarde, E. (2012). Mind wandering and driving: Responsibility case-control study. *British Medical Journal*, *345*, e8105.

Gong, J., Jiang, Y., Xiong, G., Guan, C., Tao, G., & Chen, H. (2010). The recognition and tracking of traffic lights based on color segmentation and camshift for intelligent vehicles. In *2010 IEEE Intelligent Vehicles Symposium* (pp. 431–435). New York, NY: IEEE.

He, J., Becic, E., Lee, Y. C., & McCarley, J. S. (2011). Mind wandering behind the wheel performance and oculomotor correlates. *Human Factors*, *53*, 13–21.

Hu, W., Tan, T., Wang, L., & Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *34*, 334–352.

InSight SHRP 2 NDS. (2016). Homepage. Retrieved from https://insight.shrp2nds.us/projectBackground/index

Jabon, M., Bailenson, J., Pontikakis, E., Takayama, L., & Nass, C. (2011). Facial expression analysis for predicting unsafe driving behavior. *IEEE Pervasive Computing*, *10*(4), 84–95.

Latko, W. A., Armstrong, T. J., Foulke, J. A., Herrin, G. D., Rabourn, R. A., & Ulin, S. S. (1997). Development and evaluation of an observational method for assessing repetition in hand tasks. *American Industrial Hygiene Association Journal*, *58*, 278–285.

Lee, D., Oh, S., Heo, S., & Hahn, M. (2008). Drowsy driving detection based on the driver's head movement using infrared sensors. In *Second International Symposium on Universal Communication* (pp. 231–236). New York, NY: IEEE.

Lee, J. D., Regan, M. A., & Young, K. L. (2008). Defining driver distraction. In M. A. Regan, J. D. Lee, & K. L. Young (Eds.), *Driver distraction: Theory, effects, and mitigation* (pp. 31–40). Boca Raton, FL: CRC Press.

Mather, M., Cacioppo, J. T., & Kanwisher, N. (2013). Introduction to the special section: 20 years of fMRI—what has it done for understanding cognition? *Perspectives on Psychological Science*, *8*(1), 41–43.

McAtamney, L., & Corlett, E. N. (1993). RULA: A survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics*, *24*, 91–99.

Metz, B., & Krueger, H. P. (2010). Measuring visual distraction in driving: The potential of head movement analysis. *IET Intelligent Transport Systems*, *4*, 289–297.

Mota, S., & Picard, R. W. (2003). Automated posture analysis for detecting learner's interest level. In *Conference on Computer Vision and Pattern Recognition Workshop* (Vol. 5, pp. 49–49). New York, NY: IEEE.

Mourant, R. R., & Rockwell, T. H. (1970). Mapping eye-movement patterns to the visual scene in driving: An exploratory study. *Human Factors*, *12*(1), 81–87.

National Highway Traffic Safety Administration. (2010). *Traffic safety facts: Distracted driving 2009*. Washington, DC: Author. Retrieved from http://www.distraction.gov/research/pdf-files/distracted-driving-2009.pdf

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231.

Radwin, R. G., Azari, D. P., Lindstrom, M. J., Ulin, S. S., Armstrong, T. J., & Rempel, D. (2015). A frequency-duty cycle equation for the ACGIH hand activity level. *Ergonomics*, *58*, 173–183.

Smith, B. M., Wang, X., Hu, Y. H., Dyer, C. R., Chitturi, M. V., & Lee, J. D. (2016, January). *Automatic driver face state estimation in challenging naturalistic driving videos*. Paper presented at the Transportation Research Board 95th Annual Meeting, Washington, DC.

Wang, R., Guo, L., Tong, B., & Jin, L. (2004). Monitoring mouth movement for driver fatigue or distraction with one camera. In *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems* (pp. 314–319). New York, NY: IEEE.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*, 449–455.

Yanko, M. R., & Spalek, T. M. (2013). Route familiarity breeds inattention: A driving simulator study. *Accident Analysis & Prevention*, *57*, 80–86.

Yanko, M. R., & Spalek, T. M. (2014). Driving with the wandering mind: The effect that mind-wandering has on driving performance. *Human Factors*, *56*, 260–269.

Robert G. Radwin is a professor at the University of Wisconsin–Madison in biomedical engineering, industrial and systems engineering, and orthopedics and rehabilitation. He has a BS degree (1975) from New York University Polytechnic School of Engineering, and he earned MS degrees (1979) and a PhD (1986) from the University of Michigan (Ann Arbor).

John D. Lee is the Emerson Electric Professor in Industrial and Systems Engineering at the University of Wisconsin–Madison. He graduated with a PhD in mechanical engineering from the University of Illinois in 1992.

Oguz Akkas is a graduate student research assistant at the University of Wisconsin–Madison in industrial and systems engineering.

Date received: January 7, 2016
Date accepted: February 3, 2017