Development of a Protocol to Classify Drivers' Emotional Conversation

Gregory M. Fitch[1], Richard Hanowski[1], Gary Burnett[2] and David Crundall[2]

Virginia Tech Transportation Institute[1]

University of Nottingham, UK[2]

Abstract

We present a video data reduction protocol that was developed to classify the type and intensity of emotion expressed by a driver involved in a cell phone conversation. Although there has been substantial research on coding emotion from a person's facial expression, the prescribed methods are significantly detailed and require a hundred hours of training as well as certification. The objective of the current research was to develop a simplified protocol based on the previous research that could be applied to naturalistic driving data by data reductionists in a reasonable amount of time to distill information pertaining to the nature of the driver's cell phone conversation. We present the basis for the protocol and describe how it was applied. In identifying an emotion from a driver's facial expression using naturalistic driving data, it will become possible to compute how often drivers engage in emotional cell phone conversation, the risk of a Safety-Critical Event (SCE) associated with emotional cell phone conversation, as well as the relative risk of emotional cell phone conversation compared to neutral cell phone conversation. In conclusion, we discuss other related applications of the protocol, how the protocol could be developed further and potential synergy with prior research.

Introduction

The National Highway Traffic Safety Administration (NHTSA) reports that 3,331 people were killed (10 percent of all fatal crashes), and 387,000 people were injured (17 percent of all injury crashes), in crashes involving distracted drivers in 2011 (National Highway Traffic Safety Administration, 2013). Of the 3,331 people killed in distraction-affected crashes, 385 fatalities (12%) occurred in crashes in which at least one of the drivers was using a cell phone at the time of the crash (National Highway Traffic Safety Administration, 2013). These statistics demonstrate that cell phone use plays a critical role in transportation safety.

Cell phone use comprises various subtasks, such as reaching, dialing, and conversing, that each impose differing levels of distraction. Research has unanimously found that visual-manual subtasks in which the driver removes their eyes from the forward roadway degrade driving performance and increase risk. For instance, empirical studies conducted in driving simulators, on test tracks, and with instrumented vehicles on public roads have demonstrated that visual-manual distraction degrades driving performance in terms of longitudinal vehicle control, lateral vehicle control, and response time to unexpected events (Angell et al., 2006; Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006; National

Highway Traffic Safety Administration, 2012; Neurauter, Hankey, Schalk, & Wallace, 2012; Owens, McLaughlin, & Sudweeks, 2010; Ranney, Baldwin, Parmer, Martin, & Mazzae, 2011). Naturalistic driving studies (NDSs), where drivers are continuously recorded as they operate their own vehicles without an experimenter, have demonstrated that visual-manual distraction is also associated with an increased risk of a safety-critical event (SCE; i.e., crash, near-crash, crash-relevant conflict, and unintentional lane departure) (Fitch et al., 2013; Hickman & Hanowski, 2012; Klauer et al., 2006; Klauer, Guo, Sudweeks, & Dingus, 2010; Olson, Hanowski, Hickman, & Bocanegra, 2009).

However, the impact of the cognitive demands inherent in conversing on a cell phone on safety has been debated. Some empirical research has shown that conversing on a cell phone degrades driving performance by reducing the area that drivers scan (Atchley & Dressel, 2004; Maples, DeRosier, Hoenes, Bendure, & Moore, 2008), increasing reaction time to unexpected events (Caird, Willness, Steel, & Scialfa, 2008; Horrey & Wickens, 2006), and leading drivers look at, but fail to remember seeing, billboards (Strayer, Drews, & Johnston, 2003). Drivers also report that conversing while driving increases mental workload (Horrey, Lesch, & Garabet, 2009). A study released by the AAA Foundation for Traffic Safety found that when drivers conversed on a cell phone they were less likely to scan critical locations when travelling on residential streets, and exhibited delays in responding to a red LED light mounted to a headband (Strayer et al., 2013). Despite these results, NDSs have not found conversing on a cell phone to be associated with an increased SCE risk (Hickman, Hanowski, & Bocanegra, 2010; Klauer et al., 2006; Klauer et al., 2010; Olson et al., 2009). This has been the case for both commercial motor vehicle and light vehicle drivers, as well as across broad classifications of low, moderate and high driving task demands (Fitch & Hanowski, 2011). In some cases, conversing on a hands-free cell phone was found to be associated with a decreased risk (Hickman et al., 2010; Olson et al., 2009). NDS research continues to investigate whether there are circumstances in which conversing on a cell phone increases SCE risk.

In this effort, an understanding of how often drivers converse on a cell phone is necessary to assess its population attributed risk (McEvoy & Stevenson, 2009; Sahai & Khurshid, 1996; World Health Organization, 2009). According to the National Occupant Protection Use Survey (NOPUS), drivers' hand-held cell phone use has stood around 5 percent since 2002 (National Highway Traffic Safety Administration, 2011). These data are acquired by observing drivers' electronic device use while stopped at randomly-sampled intersections across the United States. Although NOPUS provides the only probability-based observed data on driver electronic device use in the United States, it is unable to observe directly drivers' hands-free cell phone use because observers stand outside the vehicles. As such, hands-free cell phone use has been estimated by applying a hands-free to hand-held ratio calculated from the Motor Vehicle Occupant Safety Survey (MVOSS) data. From this manipulation, the National Highway Traffic Safety Administration (NHTSA) estimates that 9 percent of drivers were using either a hand-held or hands-free cell phone while driving in 2010 (National Highway Traffic Safety Administration, 2011).

As a compliment to NOPUS, NDSs have also estimated hands-free cell phone use. An analysis of the 100-Car naturalistic driving study dataset, which comprises data collected between 2003 and 2004 in the Northern Virginia and Washington D.C. area of the United States, found that drivers conversed on a cell phone 7.6 percent of the time the vehicle was travelling above 8 km/h (Fitch et al., 2013). Furthermore, a similar analysis of a commercial driver distraction dataset, which comprised data from 202 long-haul drivers between 2005 and 2007, found that professional truck drivers conversed on a cell phone 8.3 percent of the time (Fitch et al., 2013). Given how frequently drivers converse on a cell phone, there is a need to better understand the role that the inherent cognitive demands play on driver performance and crash risk. Directly measuring conversation workload in a field study, however, is difficult.

One approach has been to systematically increase the cognitive demands using the n-back task while driving on public roadways (Mehler, Reimer, & Dusek, 2011; Reimer, Mehler, Wang, & Coughlin, 2012). The n-back task involves having drivers listen to randomly ordered auditory stimuli (single digits 0–9) and responding by immediately repeating out loud either the last number heard (0-back), the second-last number heard (1-back), or the third last-number heard (2-back). Although the method serves the systematic manipulation of cognitive demands on drivers, it is not naturalistic in the sense that a researcher is seated in the rear of the vehicle for the duration of the experiment and that the task is not representative of everyday cognitive activity while driving. Much has been gleaned from the n-back task results, including that gaze concentration appears at low levels of cognitive demand, it appears prior to the appearance of marked decrements in driving control, and that it is difficult to use driving performance measures to index differences in workload prior to capacity saturation (Reimer et al., 2012). Given these results, there is a need to further assess the effect of cognitively-demanding natural conversation in a naturalistic driving study.

There have been several studies indicating how different emotions can have both positive and negative effects on driving performance (Cai & Lin, 2011; Grimm et al., 2007). Unfortunately, there has been little specific research focusing on the emotional content of conversations in the driving context. In one study, Briggs, Hole, & Land (2011) found that emotional conversation, as exhibited by drivers involved in a conversation about their fear, increased their mental workload, led to more driving errors in a driving simulator, and induced cognitive tunneling. No NDS to date has investigated the effect of emotional cell phone conversation on driver performance and associated risk of an SCE. However, what if emotional conversation could be identified from the NDS data?

Objective

We set out to develop a reduction protocol to identify emotional cell phone conversation. The protocol had to enable data reductionists to observe video footage of the driver's face and rate their emotional demeanor in a reasonable amount of time. Ideally, the protocol would be applied in a similar fashion to the Observer Ratings of Drowsiness protocol that was developed to identify instances of drowsy driving in the NDS datasets (Wiegand, McClafferty, McDonald, & Hanowski, 2009).

Constraints

The protocol was originally developed to be applied to an NDS dataset that had audio recorded of the driver’s voice. However, it was subsequently modified to be applied to an NDS dataset that did not have recorded audio. As such, the first constraint was that the protocol would need to be applied without knowing the context of the conversation. The second constraint was that the protocol would need to be applied relatively quickly. This is because NDS data reduction often involves inspecting thousands of samples under aggressive timelines and budgets. As such, the protocol was developed to capture an overall emotion by the driver, not specific facial muscle activations on a frame-by-frame basis. The third constraint was that there was only one view of the driver’s face available for reduction and it was taken from the right side of the driver’s face. Although this constraint was less severe, it may have prevented the reductionists from perceiving the symmetry of facial expressions.

Foundation for the Protocol – The Facial Action Coding System

There has been a significant amount of research on how to assess peoples’ emotional expression by analyzing images of their face. As early as 1978, Paul Ekman developed the Facial Action Coding System (FACS)(Ekman, 1978). FACS is a detailed, anatomically-based coding system that describes how to categorize facial behaviors by perceiving the activation and relaxation of specific facial muscles, called action units. The coding scheme has been updated several times with the latest version being FACS 2002. Typically, 100 hours of training is required to become FACS certified (Hager, 2003e). Furthermore, because of the work involved in identifying specific action units, a one-minute video typically takes three hours to code (Movellan, Frank, Bartlett, & Sejnowski, 2013). As such, the method is quite laborious.

The FACS manual teaches the specific action units, but does not teach what they mean (Hager, 2003a), thus ensuring that FACS coding can be objective without a rater’s biases. However, under the assumption that facial expressions have a communicative function and convey human emotion, there is a belief that certain facial expressions are associated with specific emotions (Hager, 2003c). EMFACS (Emotional FACS) was developed to use the objective scoring of FACS to identify facial expression (Ekman, Friesen, Irwin, & Rosenberg, 2003). EMFACS requires one to be able to identify the specific action units, their intensity, and their symmetry. Although EMFACS scoring is not done on a frame-by-frame basis – and can be done in one-tenth of the time of FACS - it does require FACS certification to know what action units are engaged (Hager, 2003b). Nevertheless, Hager, a student and employee of Dr. Paul Ekman, synthesizes decades of research by describing what facial expressions correspond with human emotions (Hager, 2003c). Although there are numerous types of emotions, scientific research has shown that people can reliably assign facial expression to seven categories of emotion (Hager, 2003c): Happy, Sad, Anger, Surprised, Fear, Disgust, and other emotion expressions. When applying FACS, raters also assess the intensity of each action unit on a five point scale : 1) Trace, 2) Slight, 3) Marked or Pronounced, 4) Severe or Extreme, and 5) Maximum (Hager, 2003d). Each intensity level possesses criteria that are present in the lower intensity levels.

Applying FACS

FACS was used to guide the development of the emotion reduction protocol. Coding the specific muscle activations on a frame-by-frame basis was beyond the scope of the research projects to which the protocol was to be applied. Similar to FACS, the protocol instructed reductionists how to classify the drivers' emotion into one of six categories (

Table 1). Certain emotions, such as Fear and Disgust, were combined into the other category to simplify the reduction. An "Unable to Determine" category was also added to allow reductionists to indicate when they were unable to classify an emotion. The description of the facial expressions was also simplified. It is recognized that, in doing so, the "Happy" category does not allow reductionists to directly indicate whether a driver exhibits a social smile (i.e., just the zygomaticus major is activated raising the corners of the mouth), or a Duchenne smile (i.e., the orbicularis oculi are activated squinting the eyes in addition to the zygomaticus major being activated. Note that a Duchenne smile is said to only be exhibited when true emotion is expressed (Ekman, Davidson, & Friesen, 1990)). Nevertheless, reductionists could capture true expressions of happiness in their intensity rating.

Table 1 Driver Emotion Reaction Definitions

Emotion	**Operational Definition**
Unable to Determine	Cannot tell what emotion the driver is showing
Neutral/No Emotion Shown	The driver has a straight face, does not smile or laugh, does not gesture
Happy	The driver smiles or laughs The driver gestures in excitement
Angry/Frustrated	The driver lowers/squeezes eyebrows, wrinkling forehead The driver clenches his/her teeth The driver yells (opens mouth wide with eyebrows lowered) The driver gestures in anger/frustration The driver raises his/her upper lip or tightens lips
Sad	The driver has droopy eyebrows (raises inner eyebrows, lowers outer eyebrows) The driver frowns by lowering the outer corners of his/her lips
Surprised	The driver's eyebrows raise The driver's mouth opens The driver gestures
Other	Emotional reaction that does not fit into any other category Please define in the Notes section on the Excel Log

Once reductionists rated the type of emotion expressed by a driver, they then indicated the intensity of the emotion expression. The five point rating scale used in FACS to assess the intensity of each muscle activation was converted into a four-point scale and applied to the overall emotion. Table 2 presents the intensity levels and the operational definitions used. It is recognized that the intensity levels are less selective compared to how they are applied when rating activation units. Nevertheless, they serve to identify extreme exhibitions of emotion behind the wheel.

Table 2 Driver Emotional Intensity Reduction Definitions

Intensity of Emotion	Operational Definition
Unable to Determine	Cannot tell the intensity of the emotion
Neutral/No Emotion Shown	The driver has a straight face, does not smile or laugh, does not gesture
Slight (Emotion Somewhat Shown)	The driver no longer has a straight face However, no gesturing or head movement is observed
Marked or Pronounced (Emotion Very Much Shown)	The driver no longer has a straight face The driver gestures one time in a reserved manner The driver moves his head one time
Severe (Emotion Extremely Shown)	The driver has wide eyes and a wide open mouth The driver is screaming The driver gestures wildly, or the driver moves his head frequently

It is worth noting that the intent was to develop a protocol that did utilize raters' innate ability to determine whether a driver was showing emotion. At the level of detail prescribed by the protocol, raters might classify subtle emotions as neutral. However, that was a limitation that was acceptable given that the protocol was to be used to identify high levels of emotion.

Quality Control

To apply the reduction protocol, each reductionist is required to attend a training session. Once reduction begins, each reductionist has 100 percent of their work reviewed by a senior reductionist for a minimum of 1-2 weeks. If the reductionist's accuracy is not deemed sufficient (at least 90 percent in agreement with the senior reductionist) after the initial work period, 100 percent reviews continue until accuracy is deemed sufficient. Once the reductionist's accuracy is deemed sufficient, the review percentage for that reductionist is reduced to 50 percent of reduced events. The files that are reviewed are chosen at random. Highly accurate reductionists are dropped down to 25 percent reviews and occasionally 10 percent reviews. All reductionists are required to review the senior reviewer's notes and to only make changes that they agree with. Remaining disagreements are given to another reduction specialist for resolution. As an added measure of consistency, reviewers rotate checking the work of different reductionists on a daily basis. An inter-rater test is also conducted by using twenty randomly-selected samples across all trained reductionists.

Results

This protocol was applied to a sample of cell phone conversations recorded in a recent NDS of hand-held and hands-free cell phone use (Fitch et al., 2013). The NDS, sponsored by the National Highway Traffic Safety Administration, continuously recorded 204 drivers for an average of 31 days each in 2011. Six-second segments were randomly sampled from the cell phone conversations that took place while driving above 8 km/h. Reductionists assessed the driver's emotion and emotional intensity during these 6-s samples. It was found that drivers exhibited at least marked or pronounced emotion in 3.8 percent of the cell phone conversations (35 of 921 samples). Of these 35 emotional calls, 29 were exhibitions of happiness, 5 were exhibitions of anger, and 1 was an exhibition of surprise. No emotion was classified as "other" or "unable to determine".

The protocol was also applied to 342 SCEs and 2,308 random baseline samples observed in the study (Fitch et al., 2013). Samples that exhibited at least marked or pronounced emotional cell phone conversation were scored as emotional cell phone conversation. Emotional cell phone conversation was exhibited in 7 (0.3 percent) of the 2,308 random baseline samples. It was not exhibited in any SCE. Overall, the results suggest that emotional cell phone conversation while driving, where the driver is overtly expressing emotion, is not common.

Inter-rater tests were performed using 20 samples and 16 reductionists. With respect to rating drivers' emotion, there was an average of 94 percent agreement across reductionists. Inter-rater agreement for individual samples ranged from 50 percent (1 sample), 56 percent (1 sample), 81 percent (1 sample), 94 percent (1 sample) and 100 percent (16 samples). Among the samples with some disagreement:

- The 50 percent sample had 8 "Happy" and 8 "Neutral/No Emotion"
- The 56 percent sample had 9 "Happy" and 7 "Neutral/No Emotion"
- The 81 percent sample had 3 "Happy" and 13 "Neutral/No Emotion"
- The 94 percent sample had 1 "Happy" and 15 "Neutral/No Emotion"

Among the samples with 100 percent agreement, all samples were coded "Neutral/No Emotion" by all reductionists. Note that we did not have much in the way of extreme emotions included in this preliminary test.

With respect to rating drivers' emotional intensity, there was an average of 93 percent agreement across reductionists across all samples. Inter-rater agreement for individual samples ranged from 50 percent (2 samples), 81 percent (1 sample), 94 percent (1 sample), and 100 percent (16 samples). Although, the samples that had less than 100 percent agreement on Intensity were the same samples that had less than 100 percent agreement on Emotion. This is because "Neutral/No Emotion" on Emotion was always coded as "Neutral/No Emotion" on Intensity. In contrast, any other emotion required a specific intensity. Among the samples with some disagreement:

- All samples with disagreements except for one were disagreeing between "Neutral/No Emotion" and "Slight (Emotion Somewhat Shown)", all matching up with the above listed disagreements in Emotion.
- The one sample that differed from above (the 50 percent under Emotion) had 7 "Neutral/No Emotion" (to go along with the Neutral Emotion), 8 "Slight (Emotion Somewhat Shown)" (to go along with Happy Emotion) , and 1 "Marked or Pronounced (Emotion Very Much Shown)" (to go along with Happy Emotion)
- Among the samples with 100 percent agreement, the results were the same as in the last bullet under Emotion above.

Intended Use of the Protocol

The protocol was developed to identify emotional cell phone conversation in naturalistic driving datasets. When applied, it can reveal how often emotional cell phone conversation occurs, as well as its associated SCE risk. These results can be used to estimate the population-attributed risk of emotional cell phone conversation. Although very little emotional cell phone conversation was observed in Fitch et al. (2013), and a robust estimate of the SCE risk was not possible, applying the protocol to other NDS datasets can support the investigation of emotional cell phone conversation while driving.

Further Development of Protocol – Research Agenda

Agency for Emotion

It is noted that without an audio recording to assess the context of the conversation, it is possible that reductionists may inadvertently rate the drivers' emotion expressed towards the environmental circumstances rather than towards the person to whom they are having a conversation with on the cell phone. In an attempt to control against such misclassification, the protocol requires reductionists to only rate emotion exhibited prior to a safety-critical event. This allows the identification of emotional conversation that potentially led to an SCE, rather than the identification of emotional conversation in response to an SCE. Nevertheless, it must be recognized that one still cannot tell for certain that a specific emotion arose from the conversation, rather than the environment. What is needed is research to identify behavioral markers for phone versus environment-derived emotions. To conduct such work, access would be required to actual phone conversations in NDSs.

Related to this point, the protocol could be developed in conjunction with traditional visual distraction measures to consider the medium of communication. Emotional conversations may well attract or repel attention in specific situations with a corresponding effect on distraction. For instance, with a localized sound source (e.g. dashboard speaker), an emotional conversation may be associated with certain patterns of in-vehicle glances. Research should consider how visual and emotional distraction is linked in these cases.

Driver Demeanor and Assessing Road Rage

The emotion reduction protocol can be applied to identify general driver emotional state, as well as how that emotion state may change based on environmental conditions. Instances where extreme anger has been identified could be explored to investigate the effects of road rage on driver performance. However, if identifying road rage is the objective, a separate reduction should be performed that enables reductionists to identify extreme emotion after an SCE occurs.

Automatic Data Reduction

Much work has been performed to use computer vision to automatically code facial action units (Cohn et al., 2001; Lucey et al., 2010; Ying-Li, Kanade, & Cohn, 2001). Using EMFACS, the output could be used to classify a drivers' emotion. There is tremendous merit in applying this research to naturalistic driving datasets. This would enable the efficient identification of extreme emotion behind the wheel. It would also enable a more robust assessment of how often extreme emotion is exhibited by drivers. Reliability studies would need to be conducted to compare human versus computer classifications of emotions in this context. It is worth noting that the automatic data reduction of driver drowsiness was made possible by a similar process, where a manual data reduction process was developed and later advanced into software algorithms (Sommer & Golz, 2010; Wierwille, State University, & Simulation, 1994).

Development of the Protocol for Other Research Areas

The protocol has broader appeal beyond its current use in NDSs focused on conversations. For instance, it may be developed as a tool to use in simulator studies to evaluate a driver's emotional responses. In particular, as noted by Donker, Burnett and Sharples (2011), it may be possible to consider the validity of a simulation experience based on the nature/intensity of the emotions. For instance, in an emergency braking event one would expect strong emotions of surprise and fear, rather than happiness together with laughter.

Moreover, the protocol could be expanded to consider root cause in an accident scenario (within a simulator or road trial). Visible cues within an emotion may be expressed which indicate the perceived causal factors following a crash. Such cues would consider where an individual places their visual attention (i.e. their direction of gaze) as a focus for their emotional response.

Conclusion

We report the development of an emotional cell phone conversation protocol. The data reduction protocol can be applied to recorded video of drivers observed in naturalistic driving data. The protocol was inspired by emotion research, in particular the FACS coding system. The principal merit of the protocol is that it can be applied in a reasonable amount of time to identify pronounced emotions exhibited by drivers. Although the application of this protocol to an NDS dataset found that emotional conversation is rare, the application of this protocol to other NDS datasets may enable an assessment of how frequently drivers engage in highly emotional conversation, its associated SCE risk, and its SCE risk relative to neutral cell phone conversation. With further development, the protocol will ultimately be used to shed additional insight on the SCE risk of cell phone conversation while driving.

References

Angell, L., Auflick, J., Austria, P. A., Kochhar, D., Tijerina, L., Biever, W., Diptiman, T., Hogsett, J., & Kiger, S. (2006). *Driver Workload Metrics Task 2 Final Report. Report No. DOT HS 810 635*. Washington, D.C.: National Highway Traffic Safety Administration.

Atchley, P. & Dressel, J. (2004). Conversation limits the functional field of view. *Human Factors: The Journal of the Human Factors and Ergonomics Society 46*(4), 664-673.

Briggs, G. F., Hole, G. J., & Land, M. F. (2011). Emotionally involving telephone conversations lead to driver error and visual tunnelling. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(4), 313-323. DOI: 10.1016/j.trf.2011.02.004.

Cai, H. & Lin, Y. (2011). Modeling of operators' emotion and task performance in a virtual driving environment. *International Journal of Human-Computer Studies, 69*(9), 571-586.

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis & Prevention, 40*(4), 1282-1293. 10.1016/j.aap.2008.01.009.

Cohn, J. F., Kanade, T., Moriyama, T., Ambadar, Z., Xiao, J., Gao, J., & Imamura, H. (2001). *Final Report, CIA Contract #2000-A128400-000: A Comparative Study of Alternative FACS Coding Algorithms*: Carnegie Mellon University.

Donkor, R., Burnett, G. E., & Sharples, S. (2011). *Validating driving simulators for distraction research*. Paper presented at the International Conference on Driver Distraction and Inattention.

Ekman, P., & Friesen, W. V. (1978). *The Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto: Consulting Psychologist Press.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne Smile: Emotional Expression and Brain Physiology II. *Journal of Personality and Social Psychology, 58*(2), 342-353.

Ekman, P., Friesen, W. V., Irwin, W., & Rosenberg, E. (2003). Introduction To EMFACS 2012, from http://face-and-emotion.com/dataface/facs/emfacs_intro_authors.html

Fitch, G. M. & Hanowski, R. J. (2011). The risk of a safety-critical event associated with mobile device use as a function of driving task demands. *Proceedings of the 2nd International Conference on Driver Distraction and Inattention*.

Fitch, G. M., Soccolich, S. A., Guo, F., McClafferty, J., Fang, Y., Olson, R. L., Perez, M. A., Hanowski, R. J., Hankey, J. M., & Dingus, T. A. (2013). *The Impact of Hand-held and Hands-free Cell Phone Use on Driving Performance and Safety-Critical Event Risk. (Report No. DOT HS 811 757)*. Washington, D.C.: National Highway Traffic Safety Administration.

Grimm, M., Kroschel, K., Harris, H., Nass, C., Schuller, B., Rigoll, G., & Moosmayr, T. (2007). On the Necessity and Feasibility of Detecting a Driver's Emotional State While Driving. In A. R. Paiva, R. Prada & R. Picard (Eds.), *Affective Computing and Intelligent Interaction* (Vol. 4738, pp. 126-138): Springer Berlin Heidelberg.

Hager, J. C. (2003a). Description Of Facial Action Coding System (FACS). 2012, from http://face-and-emotion.com/dataface/facs/description.jsp

Hager, J. C. (2003b). EMFACS -- Scoring For Emotion With FACS. 2012, from http://face-and-emotion.com/dataface/facs/emfacs.jsp

Hager, J. C. (2003c). Emotion And Facial Expression. 2012, from http://face-and-emotion.com/dataface/emotion/expression.jsp

Hager, J. C. (2003d). How to Read the AU Sections. 2012, from http://face-and-emotion.com/dataface/facs/manual/AU_sections.html

Hager, J. C. (2003e). NSF Report - Facial Expression Understanding. 2012, from http://www.face-and-emotion.com/dataface/nsfrept/psychology.html

Hickman, J. S. & Hanowski, R. J. (2012). An assessment of commercial motor vehicle driver distraction using naturalistic driving data. *Traffic Injury Prevention, 13*(6), 566-574. 10.1080/15389588.2012.683841.

Hickman, J. S., Hanowski, R. J., & Bocanegra, J. (2010). *Distraction in Commercial Trucks and Buses: Assessing Prevalence and Risk in Conjunction with Crashes and Near-Crashes. Report No. FMCSA-RRR-10-049*. Washington, DC: Federal Motor Carrier Safety Administration.

Horrey, W. J., Lesch, M. F., & Garabet, A. (2009). Dissociation between driving performance and drivers' subjective estimates of performance and workload in dual-task conditions. *Journal of Safety Research, 40*(1), 7-12. 10.1016/j.jsr.2008.10.011.

Horrey, W. J. & Wickens, C. D. (2006). Examining the impact of cell phone conversations on driving using meta-analytic techniques. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 48*(1), 196-205. 10.1518/001872006776412135.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data* (No. DOT-HS-810-594). Washington, DC: NHTSA.

Klauer, S. G., Guo, F., Sudweeks, J., & Dingus, T. A. (2010). *An Analysis of Driver Inattention Using a Case-Crossover Approach on 100-Car Data: Final Report No: DTNH22-00-C-07007*. Washington, D.C.: National Highway Traffic Safety Administration.

Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 94-101.

Maples, W. C., DeRosier, W., Hoenes, R., Bendure, R., & Moore, S. (2008). The effects of cell phone use on peripheral vision. *Optometry - Journal of the American Optometric Association, 79*(1), 36-42. 10.1016/j.optm.2007.04.102.

McEvoy, S. P. & Stevenson, M. R. (2009). Measuring Exposure to Driver Distraction. In M. A. Regan, J. D. Lee & K. L. Young (Eds.), *Driver Distraction: Theory, Effects, and Mitigation*. Boca Raton, Florida: Taylor & Francis Group, LLC.

Mehler, B., Reimer, B., & Dusek, J. A. (2011). *MIT AgeLab Delayed Digit Recall Task (n-back). Working Paper 2011-3B*. Cambridge, MA: Massachusetts Institute of Technology.

Movellan, J., Frank, M. S., Bartlett, M. S., & Sejnowski, T. (2013). Fully Automatic Face Detection and Expression Recognition. Retrieved March 25, 2013, from http://mplab.ucsd.edu/grants/project1/research/face-detection.html

National Highway Traffic Safety Administration. (2011). *Traffic Safety Facts Research Notes* (No. DOT HS 811 517). Washington, D.C.: National Highway Traffic Safety Administration.

National Highway Traffic Safety Administration. (2012). *Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices, Notice of Proposed Federal Guidelines. Federal Register 77:37 (24 February 2012) p. 11200*. Washington, D.C.: National Highway Traffic Safety Administration.

National Highway Traffic Safety Administration. (2013). *Traffic Safety Facts Research Notes (No. DOT HS 811 737)*. Washington, D.C.: National Highway Traffic Safety Administration.

Neurauter, M., Hankey, J., Schalk, T., & Wallace, G. (2012). Outbound Texting. *Transportation Research Record: Journal of the Transportation Research Board, 2321*(-1), 23-30. 10.3141/2321-04.

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). *Driver Distraction in Commercial Vehicle Operations: Final Report. Contract DTMC75-07-D-00006, Task Order 3*. Washington, D.C.: Federal Motor Carrier Safety Administration.

Owens, J. M., McLaughlin, S. B., & Sudweeks, J. (2010). On-Road Comparison of Driving Performance Measures When Using Handheld and Voice-Control Interfaces for Mobile Phones and Portable Music Players. *SAE International Journal of Passenger Cars - Mechanical Systems, 3*(1), 734-743.

Ranney, T. A., Baldwin, G. H. S., Parmer, E., Martin, J., & Mazzae, E. N. (2011). *Distraction Effects of Manual Number and Text Entry While Driving. Report No. DOT HS 811 510*. Washington, D.C.: National Highway Traffic Safety Administration.

Reimer, B., Mehler, B., Wang, Y., & Coughlin, J. F. (2012). A Field Study on the Impact of Variations in Short-Term Memory Demands on Drivers' Visual Attention and Driving Performance Across Three Age Groups. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. doi: 10.1177/0018720812437274.

Sahai, H. & Khurshid, A. (1996). *Statistics in Epidemiology: Methods, Techniques, and Applications*. Boca Raton, Florida: CRC Press.

Sommer, D. & Golz, M. (2010). Evaluation of PERCLOS based current fatigue monitoring technologies. *Proceedings of the IEEE Engineering Medical Biological Society*. 4456-4459.

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J., Medeiros-Ward, N., & Biondi, F. (2013). *Measuring Cognitive Distraction in the Automobile*. Washington, D.C.: AAA Foundation.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied, 9*(1), 23 - 32.

Wiegand, D. M., McClafferty, J., McDonald, S. E., & Hanowski, R. J. (2009). *Development and Evaluation of a Naturalistic Observer Rating of Drowsiness Protocol. Report No. 09-UF-004*. Blacksburg, Virginia: The Virginia Tech Transportation Institute.

Wierwille, W. W. U. S. N. H. T. S. A. O. o. C. A. R. V. P. I., State University, V. A., & Simulation, L. (1994). *Research on vehicle-based driver status/performance monitoring : development, validation, and refinement of algorithms for detection of driver drowsiness*. Washington, DC: National Highway Traffic Safety Administration.

World Health Organization. (2009). Global health risks: mortality and burden of disease attributable to selected major risks. Retrieved March 20th, 2013, from http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf

Ying-Li, T., Kanade, T., & Cohn, J. F. (2001). Recognizing action units for facial expression analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23*(2), 97-115. doi: 10.1109/34.908962.

Young, R. (2011). Driving Consistency Errors Overestimate Crash Risk from Cellular Conversation in Two Case-Crossover Studies. *Proceedings of the Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*.

Young, R. A. (2012). Cell Phone Use and Crash Risk: Evidence for Positive Bias. *Epidemiology, 23*(1), 116-118. doi: 10.1097/EDE.0b013e31823b5efc.