Technical Report UMTRI-2006-16 July 2006

Second-Generation UMTRI Coding Scheme for Classifying Driver Tasks in Distraction Studies and Application to the ACAS FOT Video Clips

SAfety VEhicles using adaptive Interface Technology (SAVE-IT Project)
Task 3C: Performance

Serge Yee, Paul Green, Lan Nguyen,
Jason Schweitzer, and Jessica Oberholtzer

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
UMTRI-2006-16		

4. Title and Subtitle	5. Report Date
Second-Generation UMTRI Coding Scheme for Classifying Driver Activities in Distraction Studies and Application to the ACAS FOT Video Clips	July, 2006
	6. Performing Organization Code accounts 049178, 049183
7. Author(s) Serge Yee, Paul Green, Lan Nguyen, Jason Schweitzer, and Jessica Oberholtzer	8. Performing Organization Report No. UMTRI-2006-16
9. Performing Organization Name and Address The University of Michigan Transportation Research Institute (UMTRI) 2901 Baxter Rd, Ann Arbor, Michigan 48109-2150 USA	10. Work Unit no. (TRAIS)
	11. Contract or Grant No. Contract DRDA 04-4274
12. Sponsoring Agency Name and Address Delphi Delco Electronic Systems One Corporate Center, M/C E110 Box 9005, Kokomo, IN 46904-9005	13. Type of Report and Period Covered 1/05-7/06
	14. Sponsoring Agency Code

15. Supplementary Notes

SAVE-IT project

16. Abstract

This report describes the development of a new coding scheme to classify potentially distracting secondary tasks performed while driving, such as eating and using a cell phone. Compared with prior schemes (Stutts et al., first-generation UMTRI scheme), the new scheme has more distinctive endpoints for tasks and subtasks, is less subjective (e.g., no "high involvement" eating), includes codes for activities absent from prior schemes (e.g., chewing gum), and more closely links subtasks to visual, auditory, cognitive, and psychomotor task demands.

The scheme has codes for 12 tasks (use a cell phone, eat/drink, smoke, chew gum, chew tobacco, groom, read, write, type, use an in-car system, internal distraction, and converse) plus codes for drowsiness. The scheme takes several factors into account, such as where the driver is looking, where the driver's head is pointed, what the driver's hands are doing, the weather, and the road surface condition. Each main task was divided into 3 to 17 subtasks (e.g., groom using tool, reach and get phone).

This scheme was used to code video clips of drivers' faces from the ACAS field operational test. In the first pass, 2,914 video clips were coded (for task, drowsiness, weather, and road) using custom UMTRI software. In the second pass, a sample of 403 distracted and 416 nondistracted clips were coded frame by frame (15,965 frames) for the subtasks performed, gaze direction, and where the head was pointed.

17. Key Words	18. Distribution Statement
Distraction, Attention, Crashes, ITS, Driving Performance, Human Factors, Ergonomics, Safety, Usability, Telematics, Tasks Analysis	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classify. (of this report)	20. Security Classify. (of this page)	21. No. of pages	22. Price
(None)	(None)	65	

Form DOT F 1700 7 (8-72) Reproduction of completed page authorized

Second-Generation UMTRI Coding Scheme for Classifying Driver Activities in Distraction Studies and Application to the ACAS FOT Video Clips

UMTRI Technical Report 2006-16
July, 2006
Serge Yee, Paul Green, Lan Nguyen, Jason Schweitzer, & Jessica Oberholtzer
University of Michigan Transportation Research Institute
Ann Arbor, Michigan
USA

1 Primary Issues

1. What are the criteria for a good coding scheme for driver tasks and subtasks?
2. How have driver tasks and subtasks been coded in previous studies?
3. What are the strengths and weaknesses of those schemes?
4. What codes should be included in a scheme to identify driver tasks and subtasks?
5. How were the ACAS video clips selected and coded in this project?
6. How could the coding schemes and coding process be improved?

2 Criteria for a Good Coding Scheme

Descriptive/ Explanatory	**Consistent (with Engineering and Human Factors Practice)**
Addresses questions posed.	Generally consistent with literature themes.
Broadly useful.	Specifically consistent with prior studies.
Tasks are unique/do not overlap.	Task structure is consistent across tasks.
Task and subtask distinctions have practical implications.	Variable and fixed tasks and subtasks are separated.
Tasks and subtasks are theoretically interesting.	Coding simultaneous activities is supported.
Scheme helps identify resources needed (visual/auditory/cognitive/psychomotor).	Tasks and subtasks have distinct endpoints.
Scheme is consistent with source accuracy.	**Practical to Use**
Scheme is complete.	Can be coded from available data.
Scheme differentiates between human and vehicle activities.	Reasonable effort to code.
Scheme yields replicable results.	Objective, not subjective.
Extensible	Unambiguous definitions.
New codes can be added.	Simultaneous tasks can be identified.
	At their finest resolution, tasks are recordable.

3 Prior Coding of Driver Tasks and Subtasks

Stutts et al. Scheme

Activity Coding

Category	Activity Name
Phone / Pager	Phone/pager not in use Dialing phone Answering ringing phone Talking/listening
Eat / Drink	Not eating or drinking Preparing to eat/drink Eating Drinking Spilled/dropped food Spilled/dropped drink
Music / Audio	Music, radio, etc. not on Music, radio, etc. on Manipulating music controls
Smoke	Not smoking Lighting cigarette, pipe, etc. Finishing smoking Smoking
Read / Write / Groom	Not reading/writing grooming Reading or writing Grooming Reading/writing and grooming
Occupant distraction	No distraction from other occ. Baby distracting Child distracting Adult distracting
Converse	Not conversing Conversing
Internal distraction	No int. event distracting driver Manipulating vehicle controls (not radio or other audio) Falling object (not food/drink) Insect distracting Pet distracting Reach/lean/look for/pick up Other internal distraction
External distraction	No ext. event distracting driver Ext. event distracting driver

Driver State Coding

Category	Description
Hands	Both hands on wheel One hand on wheel Both hands off wheel
Eyes / Head	Eyes outside of vehicle Eyes inside of vehicle
Drowsy / Aggressive	Yawning Clear anger/ aggressiveness Clear drowsiness (head jerk, eyes droop/closed)

Note: There were also codes for driving context.

First UMTRI Scheme

Driver Activity Codes

Category	Activity Name
No task	None
Use phone	Converse, handheld
	Reach for, handheld
	Dial, handheld
	Converse, hands-free
	Reach for headset, hands-free
	Unclear activity, hands-free
Eat	Eating, high involvement
	Eating, low involvement
Drink	Drinking, high involvement
	Drinking, low involvement
Converse	Conversation
In-car system	In-car system use
Smoke	Smoking, lighting
	Smoking, reaching for cigarettes, lighter, ashtray
	Smoking (active)
Groom	Grooming, high involvement
	Grooming, low involvement
Other	Other/multiple behaviors

Driver State Codes

Cat.	Activity Name	Options
Eye	Location in first frame	0=forward scene 1=left mirror or window 2=left shoulder 3= right mirror or window 4=right shoulder 5=head down IP/lap, 6=head down center stack 7=wearing sunglasses or glare 8=unknown 9=other location
	On task in first frame?	0=no 1=yes 2=unknown
	In transition	0=no, 1=yes to forward scene 2=yes, away from fwd. scene 3=yes towards & away from forward scene 4=unknown
	Time away from fwd. scene, glances 1-4	Duration of glances in tenths of seconds
Hand	Location in first frame	0=cannot see, 1=1 hand on wheel, 1 unk. 2=both hands on wheel 3=1 hand off, 1 hand unk. 4=1 hand on, 1 hand off 5=both hands off

Driving Context Codes

Category	Options
Precipitation	0=none, 1=rain, 2=snow/sleet
Road surface	0=dry, 1=wet, 2=snow covered
Seatbelt	0=yes, 1=no, 2=unknown

3 Strengths and Weaknesses of Prior Schemes

Stutts et al.		First-Generation UMTRI	
Strengths	**Weaknesses**	**Strengths**	**Weaknesses**
Good task resolution	Lacks codes for some tasks (e.g., chewing gum)	Precise gaze coding	A few parts are subjective (involvement)
Useful alphabetic shorthand	Begin and end points could be better specified	Recognizes task vary in intensity (involvement)	Lacks codes for some tasks (e.g., chewing gum)
Some linkage with CDS codes		Proven through use	Begin and end points could be better specified
Good coding of internal distractions			
Proven through use			

4 Current Coding of Driver Tasks and Subtasks

Example Eye State Coding
Looking forward at forward scene
Looking at left outside mirror or left window
Looking back over left shoulder
Looking at right outside mirror or right window
Looking back over right shoulder
Looking forward at rear-view mirror
Eyes down, looking at instrument panel or at lap area
Eyes down, looking at center stack counsel area
Transition (eyes not focused on anything)
Cannot evaluate eye location (sunglasses, glare, etc.)
Blink (eyes closed)
Other – Eyes elsewhere

Also:
11 categories for head gaze
17 categories for hands
3 categories for weather/ visibility (9 total CDS codes)
3 categories for road surface condition (7 total CDS codes)

Task Coding		
#	Task/ Category	# of Subtasks
	None	
1	Use a phone	7
2	Eat/Drink	12
3	Smoke	6
4	Chew tobacco	4
5	Chew gum	9
6	Groom	5
7	Read	3
8	Write	3
9	Type	5
10	Use in-car system	7
11	Internal distraction	5
12	Converse	6

Example Subtask Coding (3-13 subtasks typical)		
Subtask	Begin	End
Prepare to groom	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for grooming tool or to perform grooming task with hand.	Subject initiates another grooming subtask.
Groom - hand only	Subject touches grooming area with hand.	Subject removes hand from grooming area.
Groom - using tool	Subject touches grooming area with grooming tool.	Subject removes hand holding grooming tool from grooming area.
Hold grooming tool	Subject holds grooming tool in hand while not touching the grooming area.	Subject initiates another grooming subtask.
Finish grooming	Subject removes hand or grooming tool from grooming area.	Subject moves hand to a resting position or initiates another subtask.

5 Current Video Clip Selection and Coding

Pass 1: 2,914 4 s clips roughly equally chosen from 36 cells (6 road types x 3 ages x 2 sexes)
* Goal was to develop most sensitive test for road type, age and sex differences
* Clips each coded for task, drowsiness, weather/visibility, and road surface condition

Pass 2: 403 distraction / 416 nondistraction clips; sampled using the same task frequency as in Pass 1
* Goal was to focus on tasks most commonly distracting
* Clips coded by frame (20/clip, 15,965 frames) for subtask, direction of gaze, head direction, hand status
* Each clip was coded by 2 of 3 analysts in both passes in an iterative manner

PREFACE

This report is one of a series that describes the second phase of UMTRI's work on the SAVE-IT project, a federally-funded project for which Delphi serves as the prime contractor and UMTRI as a subcontractor. The overall goal of this project is to collect and analyze data relevant to distracted driving and to develop and test a workload manager. That workload manager should assess the demand of a variety of driving situations and in-vehicle tasks to determine: (1) what information should be presented to the driver (including warnings), (2) how that information should be presented, and (3) which tasks the driver should be allowed to perform. UMTRI's role is to collect and analyze the driving and task demand data that served as a basis for the workload manager, and to describe that research in a series of reports.

In the first phase, UMTRI completed literature reviews, developed equations that related some road geometry characteristics to visual demand (using visual occlusion methods), and determined the demands of reference tasks on the road and in a driving simulator.

The goals of this phase were to determine: (1) what constitutes normal driving performance, (2) where, when, and how secondary tasks occur while driving, (3) whether secondary tasks degrade driving and by how much, (4) which elements of those tasks produce the most interference, (5) how road geometry and traffic affect driving workload, (6) which tasks drivers should be able to perform while driving as a function of workload, and (7) what information a workload manager should sense and assess to determine when a driver may be overloaded.

In the first report of this phase (Yee, Green, Nguyen, Schweitzer, and Oberholtzer, 2006) (this report), UMTRI developed a second-generation scheme to code: (1) secondary driving tasks that may be distracting (eating, using a cell phone, etc.), (2) subtasks of those tasks (grooming, using a tool, etc.), (3) where drivers look while on the road, and (4) other aspects of driving. The scheme was then used to code video data consisting of face clips and forward scenes from the advanced collision avoidance system (ACAS) field operational test (FOT). The ACAS FOT was a major study in which instrumented vehicles collected a combined 100,000 miles of driving data on more than 100 drivers, who used those vehicles for everyday use (Ervin, Sayer, LeBlanc, Bogard, Mefford, Hagan, Bareket, and Winkler, 2005).

Yee, Green, Nguyen, and Schweitzer (2006) used the second-generation UMTRI coding scheme to determine how often various secondary tasks and subtasks occur as a function of the type of road driven, driver age, driver sex, and other factors. In addition, Yee, Nguyen, Green, and Oberholtzer (2006) performed an analysis to identify the visual, auditory, cognitive, and psychomotor (VACP) demands of all subtasks observed, and determined how often those subtasks were performed. The goal of this analysis was to gain insight on the degree to which various aspects of subtask demand (VACP dimensions) affect driving.

In a subsequent study, Eoh, Green, and Hegedus (2006), examine various combinations of measures (e.g., steering wheel angle and throttle) to analyze their joint distribution as a function of road type. This is done by pairing or grouping these measures to identify abnormal driving. By using the nonparametric distributions that describe these measures, pairs of thresholds were used to identify when particular maneuvers (e.g., lane change) occurred on various road types. Success in this study was truly mixed, with high detection performance in some situations, poor in others. Nonetheless, some of these thresholds were descriptive enough to be used for a preliminary workload manager.

To support a more precise description of driving, Green, Green, and Eoh (2006) developed distribution models that describe many of the driving performance measures examined.

Finally, to characterize different driving situations and tasks, Schweitzer and Green (2006) asked subjects to rate clips of scenes from the ACAS FOT data relative to 2 anchor clips of expressway driving (1 of light and 1 of heavy traffic). Scenes of expressways, urban roads, and suburban driving were used for these ratings. Subjects also identified whether or not they would manually tune a radio, dial a cell phone, or enter a navigation destination in each of the clips. This data was used to determine the probability that each of the 3 tasks would be performed on each road type as a function of rated workload. In addition, the analysts used the ACAS driving performance data to develop equations that relate workload ratings to the driving situation (e.g., amount of traffic, headway to a lead vehicle).

The next task is for Delphi to use the findings from these reports to develop a workload manager and test it.

TABLE OF CONTENTS

INTRODUCTION

One of the most thoughtful papers in human factors literature is Gould and Lewis's paper entitled "Designing for Usability: Key Principles and What Designers Think" (*Communications of the ACM*, 1985). That paper states that there are 3 key principles that designers should follow to ensure the final product or service is easy to use. Those principles are: (1) early focus on users and tasks, (2) empirical measurement, and (3) iterative design. Surprisingly, most designers do not mention these points when asked to name the major factors in designing for usability, even though these principles seem obvious,.

There has been abundant activity over the last few years concerning a particular aspect of usability, the topic of driver workload/overload/distraction (Michon, 1993; Parkes and Franzen, 1993; Kiger, Rockwell, and Tijerina 1995; Eby and Kostyniuk, 2004; Green, 2004; Neale, Dingus, Klauer, Sudweeks, and Goodman, 2005; Stutts, Feaganes, Reinfurt, Rodgman, Hamlett, Gish, and Staplin, 2005). To be consistent with Gould and Lewis's task-centric and empiric measurement principles, it is important both to identify the secondary tasks that overload/distract drivers from the primary driving task and to quantify the tasks' frequencies and durations. To identify and quantify these aspects of the secondary tasks, a coding scheme is needed.

Accordingly, this report describes the characteristics of a good coding scheme, findings from prior research, and the development and application of a new scheme. More specifically, 6 questions were addressed:

1. What are the criteria for a good coding scheme for driver tasks and subtasks?

2. How have driver tasks and subtasks been coded in previous studies?

3. What the strengths and weaknesses of those schemes?

4. What codes should be included in a scheme to identify driver tasks and subtasks?

5. How were the ACAS video clips selected and coded in this project?

6. How could the coding schemes and coding process be improved?

WHAT ARE THE CRITERIA FOR A GOOD CODING SCHEME FOR DRIVER TASKS AND SUBTASKS?

By definition, a task consists of a method and a goal. The method depends upon the object used (e.g., a handheld phone) or information manipulated (e.g., a spoken message) and the action performed (e.g., dialing). Many tasks have at least 3 phases: preparation (e.g., reaching or getting something), execution (e.g., performing a task), and completion (e.g., putting something away) (Zandin, 2003). When multiple tasks are performed concurrently, each task should be separately identified and coded. The alternative would be to note the presence of multiple activities, without specifying which ones are occurring. Tasks can be coded with a single variable (the task number) followed by a subtask number to describe the amount, type of subtask, or tool used.

The classification of human activity, tasks and subtasks, is a topic discussed by every industrial engineering textbook concerned with time and motion study (e.g., Barnes, 1980; Mundel and Darmer, 1998; Niebel and Freivalds, 2002; Salvendy, 1991). However, such studies usually focus on primarily manual activities. Therefore, with both the time and motion study literature and the human factors task analysis literature under consideration, the criteria for a good coding scheme, shown in Table 1, were developed. Determining which activities should be coded and how they should be coded is often difficult. Identifying these activities is especially challenging in this case, given the emphasis on distraction. For example, one of the central aspects of driving involves maintaining speed, which requires the driver to look at the speedometer. Is that distracting? If yes, how distracting is it? Should that task be coded as in-vehicle system use or be disregarded?

Table 1. Criteria for a Good Coding Scheme

#	Category	Criterion	Comment
1	Descriptive/ Explanatory	Immediately useful-Scheme is useful for this project (addresses questions posed).	Should be useful in deciding what does/does not interfere with driving
2		Broadly useful-Scheme is useful for addressing topics beyond the scope of this study and for other types of studies.	May be useful for non-distraction studies
3		Non-overlapping activities-The tasks, subtasks, contexts, and resources identified are unique (not overlapping).	
4		Practical implications-Task and subtask distinctions have practical implications (or driving).	Is chewing food different from chewing gum?
5		Theoretical implications-Tasks and subtasks are theoretically interesting.	
6		Resource identification-Scheme allows for distinguishing (separating) resource requirements (visual/ auditory/ cognitive/ psychomotor) and level of task and subtask demand. Within subtasks, resource demands are stable and the primary demands are consistent across instances.	Since dialing can be either voice or manual and those 2 methods lead to different resource demands, they should be different tasks.
7		Consistent with source accuracy-Scheme recognizes the limits of the data collected in terms of what can be seen and timing accuracy.	In the ACAS study, only the area at the top of the steering wheel was visible. The hands were often not visible to be coded.
8		Complete-The scheme is complete (comprehensive).	All tasks, subtasks, and contexts listed actually occur in driving.

#	Category	Criterion	Comment
9		Human vs. vehicle differences-Tasks and subtasks should differentiate between human (e.g., listening) and vehicle (e.g., speaking) activities.	
10		Replicable-Multiple video analysts should be able to take the same scheme and the same driving data, and obtain similar results.	A list of guidelines or principles may be needed to guide the application of the scheme and avoid misinterpretation of task definitions.
11	Consistent (with engineering and human factors practice)	Generally consistent-Builds upon conventions, concepts, and nomenclature in the literature, e.g., from predetermined time systems for task analysis such as MTM (Methods-Time-Measurement, (Karger and Bayha, 1987)) and MOST (Zandin, 2003).	Neither Stutts et al. nor the current ACAS scheme do this well. The idea that many tasks have 3 parts (get, execute, put) comes from MOST.
12		Specifically consistent-Similarity to specific prior studies (e.g., Stutts et al., ACAS) allows for comparison of this dataset with those sets.	The ACAS scheme is being modified slightly to code data from another study (road departure crash warning or RDCW).
13		Across task consistency-Task structure is consistent across tasks.	Most tasks can be partitioned into preparing to do them, doing them, and then ending them.
14		Separates fixed and variable-Variable tasks and subtasks are separated from fixed tasks and subtasks.	If a driver performs an activity 5 times in succession, each activity should be recorded separately, so the duration and variability of a task can be readily determined and predicted.

#	Category	Criterion	Comment
15	Practical to Use	Sufficient information is available-Task, context, and resource can be coded from available data.	Sound can be useful, and a scheme should account for sound in studies where audio data is available.
16		Distinct endpoints-Task, context, or resource is distinguishable from others; the task has distinct beginning and end points; context and resource category limits are well defined.	Video data with a higher frame rate demands more rigorous beginning and end points, and greater care with application.
17		Reasonable effort-Level of effort required to code the data is not excessive.	Recoding the existing data will not take much effort.
18		Objective-Coding is objective, not subjective (avoid high/low demand codes).	Argues against ACAS involvement codes
19		Unambiguous definitions-Description of tasks and subtasks is unambiguous so coding is repeatable.	Definitions of tasks and subtasks must be sufficiently clear that multiple analysts will code activities the same way. An example is distinguishing conversation with a passenger from conversation on a phone.
20		Simultaneity-The scheme should support the coding of simultaneous tasks.	Multiple codes can be assigned at the same time.
21		Resolvable detail-At their finest resolution, tasks and subtasks should be recordable.	Scheme definitions do not break down at smaller time scales (no gaps in activity).
22	Extensible	New codes-Scheme allows for the addition of codes.	

Several prior studies have developed or used schemes to classify driver tasks and subtasks. These schemes were reviewed and served as the basis for the coding scheme used in this project. Studies that were reviewed in detail (Stutts et al., ACAS/RDCW) are described below. The project team had hoped to include the 100-car study (Neale, Dingus, Klauer, Sudweeks, and Goodman, 2005) among the studies considered, but the relevant reports were not released in time to be included in this analysis. Barr, Yang, and Ranney (2003) was not considered because of a lack of detail pertinent to this analysis.

HOW HAVE DRIVER TASKS AND SUBTASKS BEEN CODED IN PREVIOUS STUDIES AND WHAT ARE THE STRENGTHS AND WEAKNESSES OF THOSE SCHEMES?

Stutts, et al.

To determine the differences between normal and distracted driving, Stutts, Feaganes, Rodgman, Hanlett, Meadows, Reinfurt, Gish, Mercadante, and Staplin (2003) conducted a naturalistic driving study. In that study, 3 video cameras (1 each for the driver, front seat, and forward scene) and other recording equipment were installed in 70 cars to observe drivers from 5 age groups (18-29, 30-39, 40-49, 50-59 and 60+). Drivers were recruited from Chapel Hill, North Carolina and just outside of Philadelphia, Pennsylvania.

For each subject, 3 hours of video data were recorded at 10 Hz (207 hours total). Three experienced coders independently analyzed the video data. Most of the coding took place in 2 passes. During the first pass, analysts identified the direction of the drivers' gaze (inside or outside the vehicle) and the drivers' hand positions (on or off the steering wheel). In the second pass, analysts identified all other behaviors (adverse events such as wandering in the lane, encroachments into other lanes, sudden braking, etc.). Occasionally a third pass was needed to collect data on very active drivers. After coding was completed, the analysts compared their results. About 65% to 70% of the results were the same for all analysts, and coding differences were discussed until a consensus was reached. Both the coding scheme and the coding process show considerable thought.

Their initial coding scheme was based on CDS (Crashworthiness Data System), a widely accepted federal crash database (http://www-nrd.nhtsa.dot.gov/departments/nrd-30/ncsa/CDS.html). CDS was established to examine the connection between driver distraction and crashes. CDS contains a random sample of approximately 5,000 tow-away crashes per year involving passenger cars, pickup trucks, and vans that are investigated in depth by field teams. The Stutts et al. scheme was designed to be coded using Observer Video-Pro software, and to include the distractions noted in prior analyses of crash data (Stutts, Reinfurt, Staplin, and Rodgman, 2001).

Table 2 lists the original CDS variables and the Stutts, et al. expansion of the codes. Notice that the major change between the 2 is that Stutts et al. included more detailed information about the objects manipulated and actions involved. For example, the CDS category of "other occupant" was expanded to identify whether the distraction was due to an infant, a child, or another adult. Smoking was expanded to include separate categories for lighting and for extinguishing a cigarette, pipe, etc. The "internal distraction" category was added to describe distractions due to operation of vehicle controls, reaching or looking for objects inside the vehicle, falling objects, pets, insects inside the vehicle, etc. The "other distraction" category was added and includes:

conversing with a passenger, reading, writing, and grooming. Finally, the "external distraction" category was modified. For several categories, comment fields were provided to describe the nature of distractions.

Table 2. Initial Stutts et al. Taxonomy of Driver Distractions

Original CDS Variables	Initial Revised Scheme
Outside person, object, event	External distraction (Nature of distraction specified in comment field)
Adjusting radio, cassette, CD	Music, radio, etc. on Manipulating audio controls, inserting tape or CD, etc.
Other occupant	Distracted by baby Distracted by child Distracted by adult
Moving object in vehicle Other device or object Vehicle or climate controls	Internal distraction: Manipulating vehicle/dashboard controls (not gearshift) Falling object (not food or drink) Insect distracting Pet distracting Reaching, leaning, looking for, picking up something (includes glove compartment) Other internal distraction
Eating / Drinking	Preparing to eat or drink Eating (bringing hand to mouth) Drinking (bringing hand to mouth) Spilled or dropped food Spilled or dropped drink
Using / dialing cell phone	Dialing cell phone Answering cell phone Talking/listening on cell phone
Smoking related	Lighting cigarette, pipe, etc. Extinguishing cigarette, pipe, etc. Smoking
Other distraction	Conversing with another occupant in vehicle Reading or writing Grooming

A final, more detailed scheme was later developed that includes both task identification (Table 3) and information on the driver's physical state (Table 4), as well as contextual information. In this version, objects and actions are described in even greater detail with a 2-letter code being assigned to each task. By keeping the codes short and memorable, analysts can easily recall and key in codes during analysis. Finally, information describing the driver's state was added; this information is particularly important for distraction studies. Contextual information includes: the presence of passenger(s) in the vehicle, light conditions, weather conditions, road type, traffic level, and whether the vehicle was stopped or moving.

Table 3. Final Stutts et al. Taxonomy for Driver Tasks

Category	Code	Activity Name	Modifiers/comments
Phone / Pager	PX	Phone/pager not in use	
	PD	Dialing phone	Handheld, Hands-free
	PR	Answering ringing phone	
	PP	Talking/listening	
Eating / Drinking	FX	Not eating or drinking	
	FP	Preparing to eat/drink	
	FF	Eating	
	DD	Drinking	
	FS	Spilled/dropped food	
	DS	Spilled/dropped drink	
Music / Audio	MX	Music, radio, etc. not on	
	MO	Music, radio, etc. on	Music type: CD, tape, radio, unknown
	MM	Manipulating music controls	
Smoking	SX	Not smoking	
	SL	Lighting cigarette, pipe, etc.	
	SF	Finishing smoking	
	SS	Smoking	
Reading / Writing or Grooming	RX	Not reading/writing or grooming	
	RR	Reading or writing	(Specify in Comments)
	GG	Grooming and reading/writing	
	RG	Grooming	
Occupant distraction	IZ	No distraction from other occupants	(May specify in comments)
	IB	Baby distracting	
	IC	Child distracting	
	IA	Adult distracting	
Conversing	CX	Not conversing	
	CC	Conversing	
Internal distraction	IX	No internal event distracting driver	
	IM	Manipulating vehicle controls (other than radio or other audio)	(Specify in comments)
	IF(E)	Falling object (not food or drink)	
	II(E)	Insect distracting	
	IP	Pet distracting	
	IR	Reaching/leaning/looking for/picking up	
	IO	Other internal distraction	
External distraction	EX	No external event distracting driver	
	EE	External event distracting driver	(Specify in comments)

Table 4. Driver Physical State

Category	Code	Description	Modifiers/comments
Hands	H2 H1 Hx	Both hands on steering wheel One hand on steering wheel Both hands off wheel	
Eyes / Head	EO EI	Eyes outside of vehicle Eyes inside of vehicle	
Drowsy / Aggressive	EY (E) EA (E) ED (E)	Yawning Clear anger/aggressiveness Clear drowsiness (head jerk, eyes drooping/closed)	

ACAS, RDCW, and the First-Generation UMTRI Scheme

One recent study on in-vehicle tasks is the advanced collision avoidance system (ACAS) field operational test (FOT) (Ervin, Sayer, LeBlanc, Bogard, Mefford, Hagan, Bareket, Winkler, 2005) conducted at UMTRI. The primary purpose of this study was not to examine distraction, but to examine the effectiveness of various warning systems. The ACAS FOT dataset provides a rich source of information pertaining to driver distraction, and this dataset was used to develop a coding scheme.

The ACAS FOT assessed the effect of combined adaptive cruise control (ACC) and forward crash warning (FCW) systems on real-world driving performance. Data was collected for 12 months in 2002-2003. Data collection involved a fleet of 10 model year 2002 Buick LeSabre passenger cars, each equipped with custom ACC and FCW systems. Each car was also equipped with 2 monochrome cameras (recording the forward scene and the driver's face) and additional instrumentation that recorded over 400 engineering variables (speed, steering wheel angle, etc.). Data collection began 5 minutes after the beginning of each trip (causing local roads to be underrepresented in the sample). The face video data was recorded for 4 seconds at 5 Hz at 5-minute intervals. The forward road scene data was recorded continuously at 1 Hz. The engineering variables were recorded continuously at 10 Hz. All data was fully synchronized.

A total of 96 subjects drove the test vehicles in this sample (though the full sample was larger). Equal numbers of men and women, in their 20's, 40's, and 60's, participated in the study. Fifteen of the subjects drove for 3 weeks and 81 drove for 4 weeks. The dataset used for this study consists of the data collected during the first week of testing only. This data was intended as baseline, naturalistic data for the ACAS FOT study since the ACAS system was not in operation.

A number of preliminary analyses of distraction were carried out using a scheme that was later refined and used to analyze data from the road departure crash warning (RDCW) FOT. That scheme is described in detail in Sayer, Devonshire, and Flannagan

(2005) and appears in Appendix A. The scheme has sections concerning driver tasks (Table 5), driver state (Table 6), and driving context (Table 7).

Table 5. Task Codes in First UMTRI Scheme

Category	Code	Activity Name
No Task	0	None
Use phone	10	Converse, handheld
	11	Reach for, handheld
	12	Dial, handheld
	20	Converse, hands-free
	21	Reach for headset, hands-free
	22	Unclear activity, hands-free
Eat	30	Eating, high involvement
	31	Eating, low involvement
Drink	40	Drinking, high involvement
	41	Drinking, low involvement
Converse	50	Conversation
Use in-car system	60	In-car system use
Smoke	70	Smoking, lighting
	71	Smoking, reaching for cigarettes, lighter, ashtray
	72	Smoking (active)
Groom	80	Grooming, high involvement
	81	Grooming, low involvement
Other	90	Other/multiple behaviors

Table 6. Driver State Codes in First UMTRI Scheme

Category	Activity Name	Options
Eyes	Location in first frame	0=forward scene 1=left mirror or window 2=left shoulder 3=right mirror or window 4=right shoulder 5=head down IP/lap 6=head down center stack 7=wearing sunglasses or glare 8=unknown 9=other location
	On task in first frame?	0=no 1=yes 2=unknown
	In transition	0=no 1=yes to forward scene 2=yes, away from forward scene 3=yes towards and away from forward scene 4=unknown
	Time away from forward scene, glances 1-4	Duration of glances in tenths of seconds
Hands	Location in first frame	0=cannot see 1=one hand on wheel, one unknown 2=both hands on wheel 3=one hand off, one hand unknown 4=one hand on, one hand off 5=both hands off

Table 7. Driving Context Codes

Category	Options
Precipitation	0=none, 1=rain, 2=snow/sleet
Road surface condition	0=dry, 1=wet, 2=snow covered
Seatbelt	0=yes, 1=no, 2=unknown

One major improvement over the Stutts et al. scheme is the further classification of eating, drinking, and grooming tasks by level of involvement, though this distinction is very subjective. One problem with the scheme is that it eliminates some of the resolution found in the Stutts et al. scheme. A concern with both schemes was that neither included the full sequence for all tasks (prepare, execute, complete) and both lacked the resolution necessary to separate the visual, auditory, cognitive, and psychomotor elements. These concerns led to development of an enhanced scheme that provided the desired resolution and is the subject of this report.

WHAT CODES SHOULD BE INCLUDED IN A SCHEME TO IDENTIFY DRIVER TASKS AND SUBTASKS? - THE SECOND-GENERATION UMTRI CODING SCHEME

The second-generation UMTRI coding scheme has 3 parts: task descriptions, driver state descriptions, and driving context descriptions. A full description of the scheme is included in Appendix B. There are 12 task categories (Table 8) and 1 driver status variable (drowsy). Several of these tasks (such as chewing tobacco or chewing gum) have not appeared in previous task description schemes. Each task has between 3 and 12 subtasks, which include preparation for task, action, and completion of the task. For example, "Read" includes: reaching for a book, reading the book, and putting the book away.

Table 8. Task Codes in Second UMTRI Scheme

#	Task/ Category	# of Subtasks
	None	
1	Use a phone	7
2	Eat/Drink	12
3	Smoke	6
4	Chew tobacco	4
5	Chew gum	9
6	Groom	5
7	Read	3
8	Write	3
9	Type	5
10	Use in-car system	7
11	Internal distraction	5
12	Drowsy	3
13	Converse	6

The beginning and end points of each subtask are well specified in this generation of the coding scheme. As an example, for "finish grooming," the task begins when the groom tool or hand is moved away from the grooming area and ends when the grooming tool is put down and/or the hand is in a resting state (on the steering wheel, lap, etc.)

Table 9. Grooming Subtask Beginning and End Point Definitions

Grooming Subtask	Begin	End
Prepare to groom	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for grooming tool or to perform grooming task with hand.	Subject initiates another grooming subtask.
Groom - hand only	Subject touches grooming area with hand.	Subject removes hand from grooming area.
Groom - using tool	Subject touches grooming area with grooming tool.	Subject removes hand holding grooming tool from grooming area.
Hold grooming tool	Subject holds grooming tool in hand while not touching the grooming area.	Subject initiates another grooming subtask.
Finish grooming	Subject removes hand or grooming tool from grooming area.	Subject moves hand to a resting position or initiates another subtask.

There are 3 driver state variables: 1 - eyes (12 options), 2 - head (11 options), and 3 - hands (17 options). Head direction was recorded to determine whether head orientation could serve as a surrogate for eye gaze direction, since head direction is easier to distinguish than eye gaze direction. Head and eye codes were assigned in the same manner as in the first UMTRI scheme (Table 6). However, hand code assignment was more detailed (see Appendix B) than in the first UMTRI scheme. Finally, there are 2 driving context variables: weather/visibility (3 options out of 9 CDS codes) and surface condition (3 options out of 7 CDS codes). The CDS codes were used in these situations because they were more specific than those in the first UMTRI scheme. Only 3 options each for weather/visibility and surface condition were used to examine the ACAS clips due to data limitations (low resolution and frame rate). Full utilization of the weather/visibility and surface condition CDS codes is encouraged for future studies, if the data permits this. Further details are found in Appendix B.

The development process for the UMTRI scheme was highly iterative and was based on extensive discussion of the relevant literature and design principles. Developers also met with the people coding the RDCW data. Multiple analysts coded tapes independently and results were compared. The analysts discussed differences and modified the coding scheme or the coding guidelines to reduce the likelihood of future coding differences. The analysts then recoded the clips and compared the results again. There were multiple iterations before the scheme could be consistently applied. However, the coding scheme could still be improved, such as by further clarifying the task descriptions or adding tasks.

HOW WERE THE ACAS VIDEO CLIPS PROCESSED?

How Were the ACAS Video Clips Selected for Coding?

The primary purpose of the ACAS analysis was to examine the differences between distracted and normal driving and to determine how those differences varied with road type, driver age, and driver sex (6 road types, 3 age groups x 2 sexes = 36 categories). Given this goal, data was sampled equally from each category in order to maximize the probability of detecting differences.

Table 9, below, provides definitions of the different road types of interest (ramp, interstate, freeway, arterial, minor arterial, collector, and local roads) as well as examples from the Ann Arbor, Michigan area. Notice that there are major differences in the frequency of each road type, but small differences in the relative frequency with which each road was driven between all 4 weeks and week 1. For example, 185 of the 8,951 clips (2.0%) were ramps in week 1, and 716 of the 37,416 (1.9%) were ramps over the 4 week sample. Unpaved roads were ignored for this study because there was not enough data to provide a good sample and because measurements such as lane position could sometimes not be determined.

Table 10. Road Types in the ACAS Dataset

Super-Class	Road Type	# Clips in ACAS All	# Clips in ACAS Week 1	Description	Local Examples
	Ramp	716 (1.9%)	185 (2.0%)	Roads that are connections between other roads that are not at grade or roads that connect limited access roads	ramps between M-14, US-23, and I-94
Limited access	Interstate	5022 (13.4%,)	1134 (12.7%)	A road that has limited access, crossings not at grade, and a U.S. DOT Interstate designation	I-94, I-96
	Freeway	4106 (11%)	1073 (12%)	A road that has limited access and crossings not at grade	US-23, US-12 (a 2-lane, high speed road)
Major	Arterial	1414 (3.8%)	373 (4.2%)	A primary road with access at grade and few speed changes that allow for high volume, high speed traffic movement	Plymouth Rd., Washtenaw Ave., Jackson Rd.
	Minor arterial	5443 (14.5%)	1441 (16%)	A secondary road with high volume of traffic movement, but with speed less than that of arterials; a road that connects arterials	Huron Parkway, Geddes Rd.
Minor	Collector	7174 (19.2%)	1792 (20%)	A road used to distribute traffic between neighborhoods that generally connects with arterials and limited access roads	Green Rd., Nixon Rd., Hubbard Rd.
	Local	3066 (8.2%)	834 (9.3%)	A road used to distribute traffic in and around neighborhoods	Baxter Rd.
	Unpaved	245 (0.6%)	52 (0.6%)		Warren Rd.
	Unknown	10230 (27.3%)	2067 (23%)	A parking lot or public/private facility not designated as a public road	UMTRI parking lot, Detroit metro airport lots
	TOTAL	37416	8951		

The distinction between interstate and freeway data is important and can be used to check the consistency of the data. Since the 2 road types are so similar, they should produce similar results (same mean speed, same lane variance, etc.). Although establishing consistency for these 2 road types is not sufficient to proving the consistency of the entire dataset, it is useful as a reasonable indicator. (Keep in mind that the ACAS clips are biased towards interstate and freeway driving, as the driver's face was not recorded until 5 minutes after the vehicle was started.)

In studies of driver performance, differences are often found due to age (Green, 2001) as well as age x sex interactions, and that was expected here. For example, drivers in the young and middle age groups used cruise control in about 10% of the samples in the complete ACAS dataset. However, female and male drivers in the old age group used cruise control in about 33% and 50% of the relevant samples, respectively. Though use of cruise control showed an interesting relationship with age and sex, it was not chosen as a selection criterion. Since the number of data points in each category would be too small, using cruise control use as a selection criterion would be counterproductive (6 road classes x 3 ages x 2 sexes x 2 cruise use (Y or N) = 72 categories). Therefore, the data was partitioned by 6 road types, 3 age groups, and the 2 sexes, which results in 36 unique combinations.

As explained in the next section, video clips were coded in two passes. Each pass required a different clip selection method to capture a sufficient amount of data, and the first pass required a selection method based on previous ACAS data. Preliminary estimates from the ACAS data (from the first week, baseline data) reveal that about 17% of all clips involved some sort of distraction (Appendix C). Given the project schedule, resources, and research questions, the analysis team decided to study 3,000 clips in Pass 1, about 510 of which were estimated to involve some distracting task. If all data were distributed evenly among the 36 categories there would be about 83 clips per category, 14 of which would involve distractions. Given that each clip contains about 4 seconds of video data, the analysis team would study an estimated 34 minutes of distracted driving. However, it is important to note that these are only estimates.

Using the ACAS estimates as the best available guide, the analysts had to extract a sample of 3,000 clips from the complete ACAS dataset, which contains about 37,416 total clips, 8,951 in week 1 alone, or about 1/3 of the clips from that week. To do this, each clip was assigned to one of the appropriate road type x age x sex categories and a random number was assigned to each clip. Then the clips were ordered according to their assigned random number, and the first 83 clips were selected from each category. When there were fewer than 83 clips available in a given category, more than 83 clips were extracted from other categories so that the final sample size would be 3,000 clips. Once a sample of 3,000 clips was compiled, the clips were observed to make sure they complied with the additional selection criteria shown in Table 11. Samples that did not meet the selection criteria were replaced to keep the category sizes consistent and to maintain the final sample size of 3,000 clips. However, some clips were removed from the sample in the final stages of the analysis. At this point a significant portion of the analysis had already occurred so it was not cost-effective to replace them. Ultimately, the final sample consisted of 2,914 clips.

Table 11. Selection Criteria for 3,000 Video Clips

Criteria	Reason
Use week 1 (baseline driving) data only	It is possible that in weeks 2, 3, and 4, when the ACC and FCW systems were active, those systems may have influenced the subjects' driving behaviors and driving may not be "normal."
Use only instances when the vehicle was moving at least 10 mph	Slower speeds included situations where the vehicle was stopped (and there is no driving data). To be consistent with the simulator video review experiment (Schweitzer and Green, 2006), such data should be excluded. Unfortunately, this eliminated some residential driving, which was generally low risk.
Select equally from (road type x age x sex) categories	For some of the categories (road x age x sex) there were not 83 useable clips. In such cases, clips from the other gender (the least important factor) served as substitutes.
Reject and replace clips of poor quality	Poor quality occurred because (1) lighting was poor (shadows, glare, etc.), which made the face difficult to see, (2) the forward camera was misaimed or out of focus, or (3) some of the engineering data associated with the clips (especially wiper status, road type, and traffic) was missing or was suspected to be faulty.
Reject and replace invalid trips	A trip may be deemed invalid due to equipment failure, an unrecognized driver (i.e., not one of the 96 subjects), etc.
Drop trip 1, driving from UMTRI	Trip 1 was essentially an opportunity for subjects to familiarize themselves with the vehicle. Therefore, data from trip 1 would not be useful.

After preliminary coding, however, it became apparent that distraction (as defined by the second-generation UMTRI scheme) was more common than the first ACAS study indicated. A task frequency report of this data by Yee, Green, Eoh, Nguyen and Schweitzer (2006) shows that about 45% of clips contain some kind of distracted behavior. This is much higher than the 17% estimate from the prior analysis of the ACAS data. This discrepancy is likely due to differences in the coding schemes, interpretation of terms, and the smaller sample size of the initial analysis. Because the distraction rate was almost 3 times the anticipated rate, coding all distraction clips frame by frame was determined to be impractical. Instead, approximately equal numbers of distracted (403) and non-distracted (416) clips (819 total, 15,962 frames) were selected to be coded frame by frame.

The distraction clips for the second pass were selected to have a composition that was consistent with Pass 1. Again, the Pass 2 clips were ordered and selected using random numbers, as in Pass 1. For example, if 10% of the distractions in Pass 1 involved cell phone use, 10% of the distractions on Pass 2 should involve cell phone use. Clips in Pass 2 should also be evenly distributed by road type x age group x sex category (as in Pass 1), rather than by frequency of occurrence in the real world (as in the complete ACAS dataset). Thus, the process involved sampling distraction clips balanced on the road type x age group x sex category (to maximize sensitivity to those differences) and then balanced on frequency of task occurrence in Pass 1 (to obtain the best estimates for the most common distractions).

How Was the ACAS In-Vehicle Video Coded?

Information such as driver actions and whether they were distracted (engaged in a secondary task and/or looking away from the road, depending on the situation) were determined by viewing the face camera video clips. Given the project resources, it was not feasible to develop a machine vision algorithm that would automatically classify what the drivers were doing and where they were looking. Instead, research assistants watched thousands of frames of video clips to collect the necessary data.

The ACAS data was 155 GB in size and was stored in a Microsoft Access database. The video data was 600 GB in size and was stored in separate files. The data was reviewed on a Dell Optiplex GX280 computer, which used a program developed by UMTRI to retrieve and view the data as shown in Figure 1. The controls used to play clips and assign codes in Pass 1 are shown in Figure 2, and the controls used in Pass 2 are shown in Figure 3.

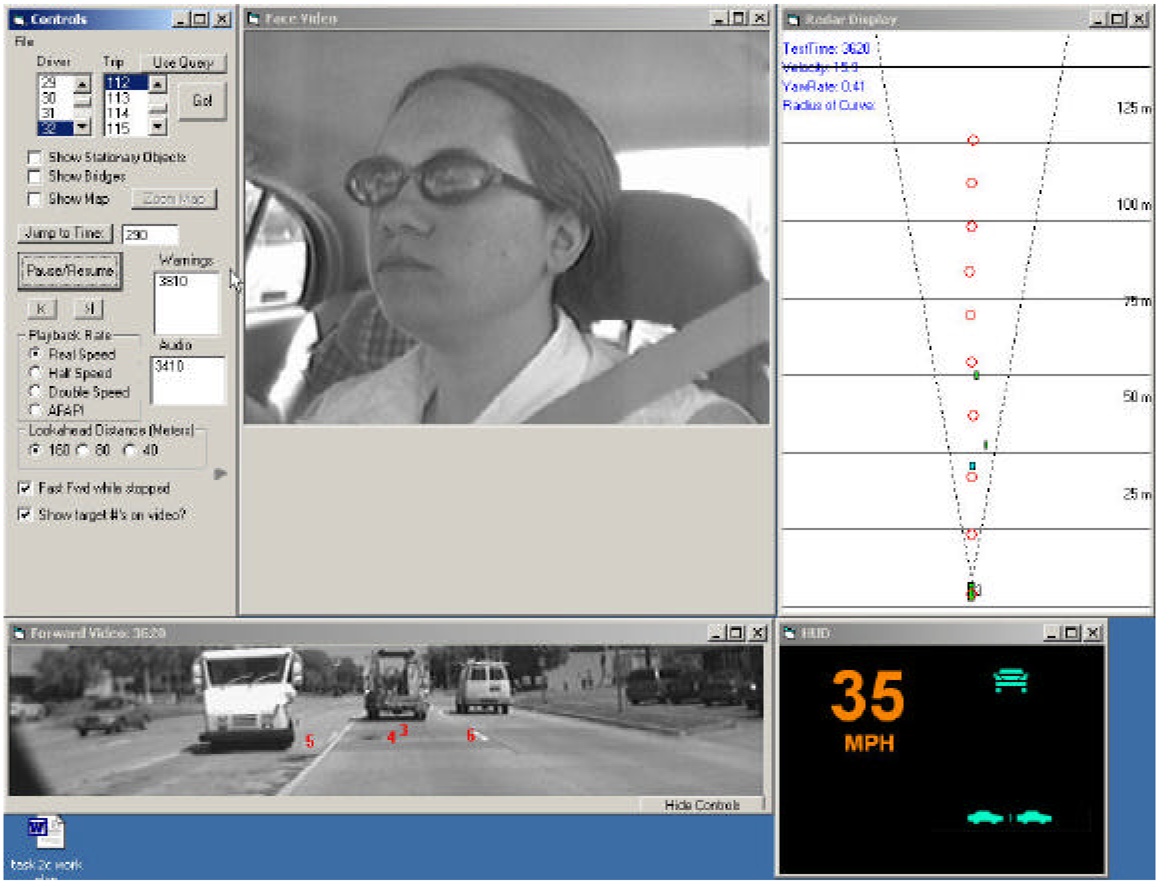

Figure 1. Screenshot of the UMTRI Software Used for Data Analysis. (Displays the driver number, clip number, playback speed, driver's face, forward radar display, speed, and forward scene)

A team of 3 analysts coded the video clips. Clips were distributed among the analysts in such a way that 2 of the analysts coded each clip. The second-generation UMTRI coding scheme (Appendix B) was used to determine: (1) where the driver was looking, (2) where the driver's head was pointed, (3) what the driver's hands were doing, and (4) which (if any) distracting activities the driver was engaged in.

The analysts initially coded the clips independently. However, when a question or discrepancy arose, the analysts would meet to confer and come to a common decision. Coding differences were usually due either to equipment and source limitations or to ambiguities in the coding scheme. Equipment and source limitations included the single camera view, the lack of audio data, and the 5 Hz frame rate. Coding scheme ambiguities included unspecific or overlapping definitions and definitions for multiple, undistinguishable distractions. These ambiguities, their resolution and the evolution of the coding scheme are summarized in the next section. In addition, coding guidelines were developed (Appendix D) to foster the development of this (and potentially future) coding scheme(s).

After the initial clip selection, a pilot test was conducted to determine how the data could be reliably coded. In this test, 2 analysts independently coded 20 clips using the preliminary second-generation UMTRI coding scheme. The test revealed important coding difficulties, such as the potential for chewing and conversation activities to be confused in the absence of audio data. Inconsistencies in the interpretation of eye glance codes also became evident. These and other challenges led to changes in the coding scheme. The pilot test also revealed that it would be best to code the clips in 2 passes. During the first pass, analysts were to identify whether any distracting tasks occurred at any time within a given clip. During the second pass, analysts coded more detailed information about the distracting task (subtask name, subtask duration, eye and head orientation, and hand position) for each frame of the clip.

Pass 1 Coding

Pass 1 coding was intended to separate clips that involve a distracting activity (for more detailed coding in Pass 2) from clips that did not. To that end, an initial assessment of the distraction rate was needed. During Pass 1, each 4-second clip was examined to determine whether a task was present; task duration was not recorded. Only the distracting task was recorded (e.g., “2: eat/drink”), not the subtask (e.g., “2.1: prepare to eat” or “2.5: chew food”). In addition, other context information was noted. (See Figure 2 for a screen shot.) In addition to helping the design of Pass 2, the experience from Pass 1 provided an opportunity to improve the second-generation UMTRI coding scheme.

Ideally, the pilot test would have provided enough information so that the clips would have had to be viewed only once during Pass 1. However, with each repetition, additional issues with the UMTRI scheme were discovered. One revision of the coding scheme was the addition of tasks. For example, “chewing gum” was added after analysis began, as chewing is believed to have an association with mental workload. In addition to obvious gum chewing activity, the task definition includes other miscellaneous mouth movements similar to gum chewing to account for the fact that gum may not always be visible. Drowsiness was observed in some of the video data, and though it is not a “task,” drowsiness was shown to have a considerable effect on driving behavior and is thus relevant to this report.

Changes to code applications constitute another revision to the UMTRI scheme. For example, during Pass 1 coding it was discovered that reading material and, occasionally, the driver’s hands could sometimes be seen as a reflection in the driver’s sunglasses. This observation was useful for detecting behaviors that would normally be out of the camera’s view. These changes to the coding scheme necessitated recoding the clips.

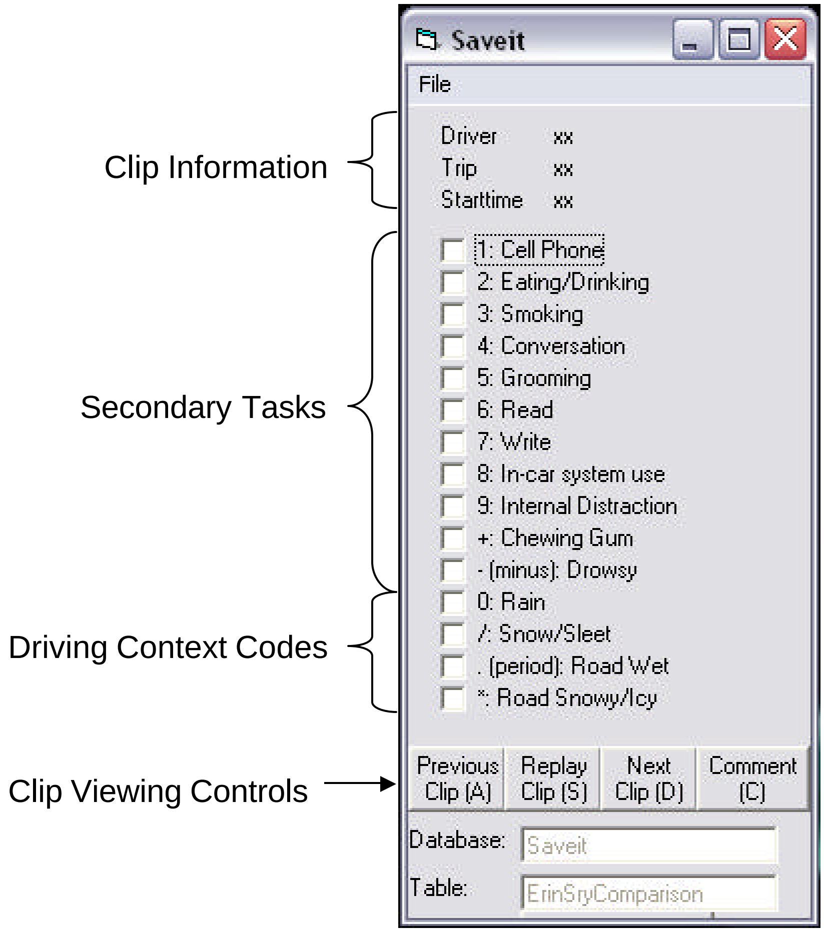

Figure 2. Clip Viewing and Rating Controls, First Pass

Pass 2 Coding

The clip viewing procedure used in Pass 1 was adapted to a higher level of detail in Pass 2. For example, in Pass 1 an analyst would simply note that grooming occurred at some time during the 4-second clip. In Pass 2, the analyst would record which grooming subtask was performed (e.g., “6.1: Prepare to groom” or “6.3: Groom – using tool”) in each frame (images were recorded at 5 Hz, so each clip could contain between 1-20 frames of grooming). In addition to subtask, the glance location, head direction, and hand positions were coded for each frame. Figure 3 shows the controls used by the video analysts for playing and coding clips.

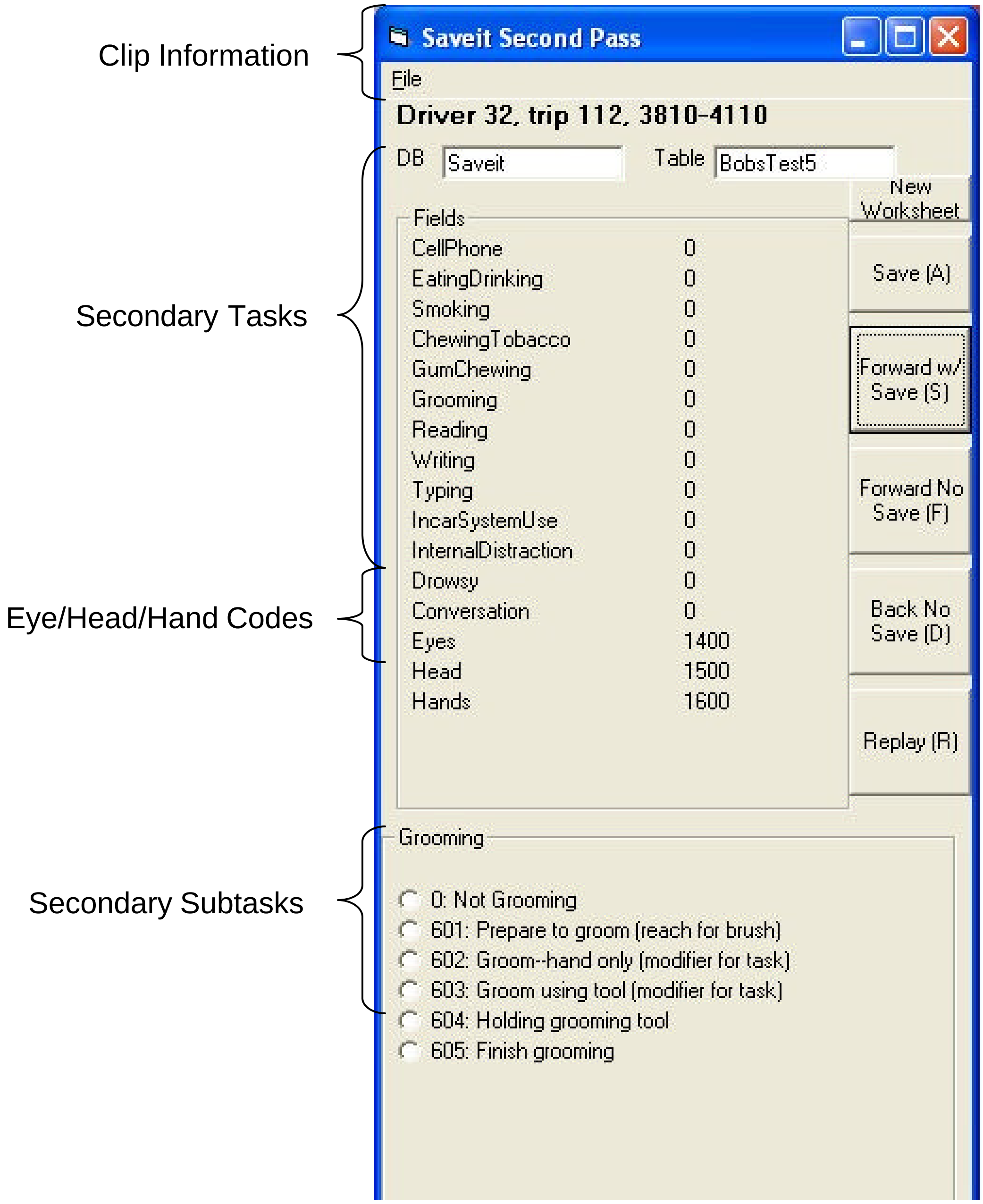

Figure 3. Clip Viewing and Rating Controls, Second Pass

To establish a suitable approach to frame-by-frame coding, a pilot test was conducted before Pass 2 began. In this pilot test, the first 20 clips containing "conversation" and the first 10 clips containing "grooming" were coded, since these tasks were expected to be the most difficult to code. As before, the analysts coded them independently and then met to compare their results and modify the coding scheme as needed.

Many of the modifications to the UMTRI scheme during the Pass 2 pilot test were applied to hand position definitions. Initially, many of the hand position-related categories specified the type of action (e.g. holding a cigarette versus holding food). To simplify coding and avoid redundancy, the codes for hand position were altered to note that the hand was performing an action, leaving the subtask codes to specify the activity (e.g., "3.5: Hold cigar/cigarette" or "1.5: Hold cell phone"). Additional codes were added to account for the driver's hand resting in a static position other than on the steering wheel.

Many of the subtask beginning and ending conditions were changed to enable analysts to more accurately and consistently describe subtask duration. For example, the initial "begin" definition for "1.6: Hang up/end call" was "Starts with the End button being pushed." The video resolution and 5 Hz frame rate made such an observation impossible, though it was relatively easy to observe the phone being removed from the ear. Furthermore, a person with a flip phone (a phone that folds open and shut) may simply fold their phones shut when ending a call without pressing End; such a motion would require that the phone is first moved away from the ear. Therefore, the "begin" definition was amended to say "Subject takes phone from ear or presses End button." Also, in some cases 2 categories that were difficult or impossible to distinguish, given the available data, were combined, such as "vertical chewing" and "horizontal chewing."

After the Pass 2 pilot study, task "4: Chewing tobacco" was added. This task differs from task "2: Eat/Drink" in that chewed material is not swallowed, and it differs from task "5: Chewing gum" in that no spitting is required during gum chewing activities. Task "4: Chewing tobacco" was not observed in any of the ACAS FOT video clips but was added to the coding scheme for completeness and possible future use.

Other changes were relatively minor. Subtask "13.3: Converse with passenger – Listen" was added to the task "13: Conversation." Task "9: Type, type on laptop" was changed to subtask "9.4: Type on full keyboard," since typing on a full keyboard is not the sole domain of a laptop computer and could be applicable to another device such as a PDA. Subtasks "10.6: Glance only – Monitor in-car systems" and "11.4: Glance only – Monitor internal distraction" were added, since glancing away from the road constitutes a distraction and because the majority of in-car system use and internal distraction subtasks were indicated by glance. The theoretical implication is that a task requiring use of the hands differs from a task requiring just a glance. This general distinction was often reflected in the report concerning visual, auditory, cognitive, and psychomotor demands, which is also part of this phase of the SAVE-IT project (Yee, Nguyen, Green, and Oberholtzer, 2006).

CONCLUSIONS

1. What are the criteria for a good coding scheme for driver tasks and subtasks?

Coding schemes should be descriptive/explanatory, consistent with engineering and human factors practice, practical to use, and extensible. Descriptive, clearly defined codes should be immediately useful for addressing questions and assigning tasks and subtasks to non-overlapping categories. They should allow for distinctions of both practical and theoretical interest, and for identification of the mental resources (visual/auditory/cognitive/psychomotor) each subtask requires. Descriptive schemes should also be replicable, such that multiple analysts can obtain similar results using the same scheme and video data. Coding schemes should also be consistent with engineering and human factors practice, existing literature (e.g., predetermined time systems), and related prior studies. They should completely segregate human and vehicle tasks as well as fixed and variable tasks (tasks done once versus multiple times in succession). Schemes that are practical to use (1) allow analysts to extract necessary data from available data, (2) have tasks and subtasks with distinct endpoints, and (3) allow each simultaneously occurring task to be coded. Such schemes are objective, unambiguously defined, and operate at the level of detail an analyst can observe and record. Extensible systems are constructed so that additional codes can be readily added.

2. How have driver tasks and subtasks been coded in previous studies?

Two previous schemes were examined, Stutts et al. and the first-generation UMTRI scheme. The Stutts et al. scheme was based on the coding scheme used by the Crashworthiness Data System (CDS) to connect driver task performance with crashes. The Stutts et al. scheme expands the 8 categories in CDS (e.g. other occupant, eating/drinking, etc.) into 9 categories with 36 total activities (e.g., distracted by baby, preparing to eat or drink, etc.). Stutts et al. also provides codes for the physical state of the driver (e.g. hands position, eyes/head direction, drowsiness/aggression) and for driving context (e.g. weather, traffic, etc.).

The first-generation UMTRI scheme (used in the ACAS and RDCW projects) uses many of the same categories as Stutts et al. (e.g. phone, eat, drink, converse, in-car system use, smoke, groom, other) as well as codes for eye direction, hand position, and driving context. The first-generation UMTRI scheme provided much greater detail on where the driver was looking than the Stutts et al. scheme and added the low/high involvement distinction for eating, drinking, and grooming. Coding for some categories, such as internal distraction, was less detailed and distinctions were handled with comments.

3. What are the strengths and weaknesses of those schemes?

Both schemes have their advantages and disadvantages. The main advantage of the Stutts et al. scheme is the emphasis on linking observed activities to codes in CDS.

The main advantage of the first-generation UMTRI scheme is the emphasis on detailed information regarding eye direction and hand position.

The main disadvantage of both schemes is the lack of a consistent coding structure across all tasks (e.g. preparation, execution, and completion), the lack of detail in defining some subtasks and their endpoints, the narrow spectrum of tasks (e.g., did not include chewing gum, chewing tobacco, or typing), and the failure to link subtasks to specific visual, auditory, cognitive, and psychomotor demands. Recognition of these disadvantages led to the creation of the second-generation UMTRI coding scheme, which builds upon these 2 preexisting schemes.

4. What codes should be included in a scheme to identify driver tasks and subtasks?

The second-generation UMTRI scheme has 12 main tasks: use cell phone, eat/drink, smoke cigar/cigarette, chew tobacco, chew gum, groom, read, write, type, use in-car system, internal distraction, and converse. Within each task there are 3-12 subtasks (e.g. “8.1: Prepare to write,” “8.2: Write,” and “8.3: Put away writing materials”). There are 3 driver state variables: eyes (12 options), head (11 options), and hands (17 options), and a code for drowsiness. The second-generation UMTRI scheme also employed 3 of the 9 CDS weather/visibility codes and 3 of the 7 CDS surface condition codes.

5. How were the ACAS video clips selected and coded in this project?

Clips were selected and coded in 2 passes. During Pass 1, clips were selected so that there would be a roughly equal number in 36 categories (6 road types x 3 ages x 2 sexes = 36 categories). For some categories, there were not enough clips in the ACAS FOT sample to provide the 83 clips needed to fill the category. When this was the case, clips from the other sex were substituted. If a clip was found to be defective during Pass 1 (poor video quality, missing engineering data, etc.), it was replaced so that the sample at the end of Pass 1 coding was 3,000 clips in size. However, if a clip was found to be defective for any reason during Pass 2 coding, it was not replaced. At this point in the coding process it was no longer feasible to replace and recode clips. After Pass 2 coding, the sample consisted of 2,914 clips.

For coding in both passes, 2 of the 3 trained analysts viewed each clip using custom UMTRI software. This software showed the driver’s face, a radar display of the traffic ahead, the forward scene, and other information. After independently reviewing each clip, the analysts met and compared the codes for each 4-second video clip. The analysts discussed differences, resolved inconsistencies, assigned codes, and in many cases, revised subtask definitions or coding guidelines.

The purpose of Pass 1 coding was to determine the overall activity of the driver and overall driving conditions. When applicable, analysts coded which distracting tasks the driver performed, driver drowsiness, and driving context variables (e.g. weather/visibility and surface condition).

It was impractical to code each clip on a frame-by-frame basis since the sample of 2,914 clips with about 20 frames each would yield about 58,280 frames. Even coding all clips that contained a distracting task (as determined in Pass 1) was impractical as the actual proportion of clips that contained a distracting task was much higher than the original estimate. Therefore, an almost equal number of distracted (403) and non-distracted (416) clips were coded in Pass 2, for a total of 15,965 frames. Distracted clips were selected so that the second pass' task frequency and category distribution remained similar to the first pass.

The purpose of Pass 2 coding was to determine which (if any) subtask is performed in each frame and the direction of gaze (where the eyes were pointed as well as head orientation). By separating the coding process into 2 passes, analysis provided both the broad overview and detailed analyses needed, and was more efficient than performing detailed coding of all clips in a single pass.

6. How could the coding schemes and coding process be improved?

The second-generation UMTRI coding scheme is an improvement upon previous schemes, but is not perfect itself. Review of the data collected (see the Preface for the full list of reports) may reveal tasks and/or subtasks that should be clarified or added. In this and several other projects, the full instrument panel was not visible, and therefore all in-car system use tasks were all in 1 category. However, had there been a good view of the instrument panel, having separate tasks or subtasks for each specific system (entertainment, climate, etc.) could make sense.

Although this coding scheme advances the concept of tasks having specific stages (preparation, execution, completion), structuring tasks as performed actions (the verb) using an object/tool (the noun) could be improved. At the time of this report, no tests have been performed to determine whether people outside the research team that developed the coding scheme will find the scheme useable or can use it consistently.

One major problem with this and all prior related work is the coding scheme's reliance on accurate manual data reduction. Ultimately, accuracy of any results based upon this coding scheme depends not only upon the competence and effort of the analyst(s), but also on the clarity and integrity of the raw video data being analyzed. In this study, for example, much of the ambiguity surrounding "conversation" could have been resolved with audio data alone. Furthermore, additional cameras (recording the driver's hands, the driver interface, and the other passengers) would have provided more information about hand position and internal distractions. Much more work needs to be done to implement data collection systems with higher-resolution video recordings (which would have enabled fuller utilization of the CDS driving context codes) and to develop tools to automate the analysis (e.g. automatically determining head and eye gaze direction, where gaze is fixated, and, most difficult, what the drivers are doing). Without those capabilities, vast quantities of data may be collected, but only a small fraction will be analyzed. In the case of the ACAS FOT, there were 34,416 clips, of which 2,914 (less than 10%) were coded for the first pass and only after a very substantial effort.

REFERENCES

Barnes, R.M. (1980). Motion and Time Study: Design and Measurement of Work (7th ed.), New York, NY: Wiley.

Barr, L. C., Yang, D. C. Y. and Ranney, T.A. (2003). Exploratory Analysis of Truck Driver Distraction Using Naturalistic Driving Data, paper presented at the 82nd Annual Meeting of the Transportation Research Board (document P03-2897. January 12-16, 2003), Washington, DC.: Transportation Research Board.

Eby, D.W. and Kostyniuk, L.P. (2004a). Crashes and Driver Distraction: A Review of Databases, Crash Scenarios, and Distracted Driving Scenarios (SAVE-IT project technical report, Task 1), Washington, D.C.: U.S. Department of Transportation (http://www.volpe.dot.gov/opsad/saveit/docs/dec04/litrev_1.pdf).

Eoh, H., Green, P.A., and Hegedus, E. (2006). Driving Performance Analysis of the ACAS FOT Data and Recommendations for a Driving Workload Manager (Technical Report UMTRI-2006-18), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Ervin, R., Sayer, J., LeBlanc, D., Bogard, S., Mefford, M., Hagan, M., Bareket, Z., and Winkler, C. (2005). Automotive Collision Avoidance System Field Operational Test Methodology and Results (Unpublished technical report), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Goodman, M, Bents, F.D., Tijerina, L., Weirwille, W., Lerner, N., and Benel, D., (1997). An Investigation of the Safety Implications of Wireless Communications in Vehicles, (Technical Report DOT HS 808-635), Washington, D.C.: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Gould, J.D. and Lewis, C. (March, 1985). Designing for Usability: Key Principles and What Designers Think, Communications of the ACM, 28(3), 300-311.

Green, P. (2001). Variations in Task Performance Between Younger and Older Drivers: UMTRI Research on Telematics, paper presented at the Association for the Advancement of Automotive Medicine Conference on Aging and Driving, Southfield, MI (http://www.umich.edu/~driving/publications.html).

Green, P. (2004). Driver Distraction, Telematics Design, and Workload Managers: Safety Issues and Solutions (SAE paper 2004-21-0022), Proceedings of the 2004 International Congress on Transportation Electronics (Convergence 2004, SAE publication P-387), 165-180. Warrendale, PA: Society of Automotive Engineers

Green, P.E., Green, P.A., and Eoh, H. (2006). How Do Distracted and Normal Driving Differ: An Analysis of the ACAS FOT Data (Technical Report UMTRI-2006-**), Ann

Arbor, MI: University of Michigan Transportation Research Institute.

Horrey, W.J. & Wickens, C.D. (2004). Cell phones and Driving Performance: A Meta-analysis, Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, 2304-2308, Santa Monica, CA: Human Factors and Ergonomics Society.

Karger, D. and Bayha, F. (1987). Engineered Work Measurement (4th ed.), New York: Industrial Press.

Kiger, S. M., Rockwell, T. H., and Tijerina, L. (1995). Developing Baseline Data on Heavy Vehicle Driver Workload, Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting, 2, 1112-1116, Santa Monica, CA: Human Factors and Ergonomics Society.

McEvoy, S.P., Stevenson, M.R., McCartt, A.T., Woodward, M., Haworth, C., Palamara, P., and Cercarelli, R., (2005). Role of Mobile Phones in Motor Vehicle Crashes Resulting in Hospital Attendance: a Case-crossover Study, British Medical Journal, August, 20, 2005, 331, 428-430.

Michon, J.A. (ed.) (1993). Generic Intelligent Driver Support. London, UK: Taylor and Francis.

Mundel, M.E. and Darner, D.L (1998). Motion and Time Study: Improving Productivity (7th ed.), Englewood Cliffs, N.J.: Prentice Hall.

Neale, V.L., Dingus, T.A., Klauer, S.G., Sudweeks, J., and Goodman, M. (2005). An Overview of the 100-Car Naturalistic Study and Findings (paper 05-0400), Experimental Safety Vehicle Conference, www-nrd.nhtsa.dot.gov/ pdf/nrd-12/100Car_ESV05summary.pdf.

Niebel, B.W. and Freivalds, A (2002). Methods, Standards, and Work Design (11th ed), New York, NY: McGraw Hill. (new edition due July 26, 2002)

Parkes, A.M., and Franzen, S. (1993). Driving Future Vehicles. London, U.K., Taylor and Francis.

Redelmeier D.A. and Tibshirani. R.J,. (1997). Association Between Cellular-Telephone Calls and Motor-Vehicle Collisions, The New England Journal of Medicine, February, 336(7), 454-458.

Salvendy, G. (ed.) (1991). Handbook of Industrial Engineering (2nd ed.), New York: Wiley.

Sayer, J.R., Devonshire, J.M., and Flannagan, C.A. (2005). The Effects of Secondary Tasks on Naturalistic Driving Performance (Technical Report UMTRI-2005-29), Ann Arbor, Michigan: University of Michigan Transportation Research Institute.

Schweitzer, J. and Green, P.A. (2006). Task Acceptability and Workload of Driving Urban Roads, Highways, and Expressway: Ratings from Video Clips (Technical Report UMTRI-2006-19), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Shelton, L.R. (2001). Statement before the Subcommittee on Highways and Transit, Committee on Transportation and Infrastructure, U.S. House of Representatives, May 9, 2001, http://www.nhtsa.dot.gov/nhtsa/announce/testimony/distractiontestimony.html.

Strayer, D. L., Cooper, J. M., & Drews, F. A. (2004). What Do Drivers Fail to See When Conversing on a Cell Phone? Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society, 2213-2217, Santa Monica, CA: Human Factors and Ergonomics Society.

Stutts, J.C., Feaganes, J.R. Reinfurt, D., Rodgman, E., Hamlett, C., Gish, K., and Staplin, L. (2005). Driver's Exposure to Distractions in their Natural Driving Environment, Accident Analysis and Prevention. 37(6), 1093-1101.

Stutts, J.C., Feaganes, J.R., Rodgman, J.R, Hanlett, E.A., Reinfurt, C., Gish, D.W., Mercadante, M., and Staplin, L. (2003). The Causes and Consequences of Distraction in Everyday Driving, Proceedings, 47th Annual Meeting, Association for the Advancement of Automotive Medicine. Des Plaines, IL.: Association for the Advancement of Automotive Medicine, 235-51.

U.S. Department of Transportation (2004). Crashworthiness Data System (CDS), National Accident Sampling System (NASS). Crashworthiness Data System (CDS) Data Set Codebook (UMTRI Report 2004-8), Ann Arbor, Michigan, University of Michigan Transportation Research Institute, (http://www.umtri.umich.edu/tdc/codebook.html).

Wang, J.-S., Knipling, R. R., and Goodman, M. J. (1996). The Role of Driver Inattention in Crashes: New Statistics from the 1995 Crashworthiness Data System, Proceedings of the Association for the Advancement of Automotive Medicine 40th Annual Conference, Des Plaines, Iowa, Association for the Advancement of Automotive Medicine, 377-392.

Yee, S., Green, P.A., Nguyen, L., and Schweitzer, J. (2006). Frequency of Distracting Tasks People Do While Driving: An Analysis of the ACAS FOT Data (Technical Report UMTRI-2006-17), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Yee, S., Green, P.A., Nguyen, L., Schweitzer, J., and Oberholtzer, J. (2006). Second-Generation UMTRI Scheme for Classifying Driver Activities in Distraction Studies and Coding ACAS Video Clips (Technical Report UMTRI-2006-16), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Yee, S., Nguyen, L., Green, P.A., and Oberholtzer, J. (2006). The Visual, Auditory, Cognitive, and Psychomotor Demands of Real In-Vehicle Tasks (Technical Report UMTRI-2006-20), Ann Arbor, MI: University of Michigan Transportation Research Institute.

Zandin, K.B., (2003). MOST Work Measurement Systems (3rd ed.), London, U.K.: Taylor and Francis (CRC Press), formerly Marcel Dekker.

APPENDIX A – FIRST UMTRI CODING SCHEME

Precipitation
Precipitation was identified via the forward camera scene, although it was sometimes difficult to know whether a given case of precipitation was rain or snow.
0 = None
1 = Rain
2 = Snow/Sleet

Road Condition
The category was used to identify whether the road was dry, wet, or covered with snow. Cues came from the forward camera scene, and included reflections on the road, precipitation, windshield wiper state, etc.
0 = Dry
1 = Wet
2 = Snow covered

Seatbelt
From the face camera scene, the driver's seat belt could usually be seen. However, because the image was black and white, the seatbelt could potentially blend into the color of the driver's clothes, making a determination difficult.
0 = Yes
1 = No
2 = Cannot tell

Location of Eyes at First Frame
Eye location was coded by what the reviewers could see of the driver's eyes at the first frame. The reviewers coded the location of the driver's eyes even if they could only see one eye, as it was assumed that the driver's eyes moved in parallel. The reviewers needed to be very confident in location of the driver's eyes in order to code as a specific location. There were many instances when the reviewers were confident that the driver's eyes were not looking forward, but could not tell specifically where the eyes were looking. The determination of whether glances were still forward or if they were glances away was also very difficult and subjective. The reviewers agreed upon an area or "box" which they considered to be looking forward. This allowed for slight glances out of the box, but multiple scans across the forward scene were considered glances away. This process defined "looking forward" very narrowly and essentially meant straight forward. Glances toward the right of the forward scene and to the right area of the windshield were considered glances away and were coded as 8.

0 = Looking forward at forward scene
1 = Left outside mirror or window
2 = Looking over left shoulder (The driver's gaze needed to look over the driver's shoulder, though the driver's chin did not necessarily need to cross over the driver's shoulder.)

3 = Right outside mirror or window
4 = Looking over right shoulder (The driver's gaze needed to look over the driver's shoulder, though the driver's chin did not necessarily need to cross over the driver's shoulder.)
5 = Head down, looking at instrument panel or lap area
6 = Head down, looking at center stack console area (the area where the entertainment system, climate control, and clock are located)
7 = Driver wearing sunglasses or glasses with glare (The glare prohibited the ability to classify where the eyes are looking. There were instances where drivers were wearing sunglasses but the reviewers felt that they could confidently identify the location of the drivers' eyes. In these instances eye location was recorded.)
8 = Cannot accurately evaluate eye location (An 8 is chosen when the reviewer was unsure of the eye position and/or classification within a reasonable level of confidence though not because of glasses. Typically the reviewer could see the actual eye, but could not determine where the gaze was directed. Eyes in transition were often coded as 8, as it was unclear where the driver's gaze was at that particular moment.)
9 = Other (For example the driver may clearly be looking at the passenger side floor. When a glance was coded as other, the location was noted in the notes section. The most common position recorded as other was the rearview mirror.)

Eyes on Task at First Frame

Not used for analyses in this report. This category defined whether the driver could be said to be paying active attention to the driving task, evidenced by his/her gaze being directed either toward the forward scene, mirrors, instrument panel, etc.

0 = No (The classification of no was only used when the reviewer could confidently determine that the driver's eyes were off the task of driving.)
1= Yes (The classification of yes does not mean looking forward, it means that the driver's eyes were on the task of driving.)
2 = Cannot determine (For instance, the driver was wearing glasses with glare or the reviewer could not see the driver's eyes for some other reason. This classification was also used when the reviewer could not tell if the eye location was on task. For instance, the driver was looking out the window but it was unclear whether the driver was looking at traffic or at a fancy building that was distracting the driver's attention. In any case, the reviewer did not KNOW whether the driver was on task or not.)

Hand Location at First Frame

Not used for analyses in this report. Both hands were not often visible, so the reviewer coded what could confidently be inferred from the scene. At times, playing the video farther helped to determine what was ambiguous in a still frame. For instance, at the first frame there may have been a small blur near the steering wheel. Upon continuation of the video the blur may have moved and come into view as a hand.

0 = Cannot see the position of either hand or cannot determine the position of either hand (The reviewer coded 0 if a hand could be seen but the reviewer could not tell if it was on the wheel.)
1 = At least one hand on steering wheel (This was coded when the position of one hand

could not be determined but one could see that at least one hand was on the steering wheel.)
2 = Both hands are on the steering wheel.
3 = At least one hand is off the steering wheel (This was coded when the position of one hand could not be determined but at least one hand was clearly off the steering wheel.)
4 = One hand on, one hand off the steering wheel. (A 4 was classified when the reviewer could clearly see both hands, and one was on the wheel while the other was off.)
5 = Both hands off the steering wheel. (A 5 was classified when the reviewer could clearly see both hands, and both were off of the wheel.)

Eyes in Transition
Not used for the analyses in this report. This category refers to instances in which the first frame of video included a transition in the driver's gaze from one direction to another.
0 = No
1 = Yes, towards forward scene
2 = Yes, away from forward scene
3 = Yes, both towards and away from forward scene
4 = Cannot tell (4 was selected when the driver was wearing sunglasses or the reviewer could not see the driver's eyes for some other reason; therefore it was uncertain whether they were in transition.)

Time Away from Forward Scene, Glances 1-4
The duration of up to four glances away from the forward scene were coded in tenths of seconds. The forward scene was defined in the same manner as for Location of Eyes at First Frame (above). If a driver was in the process of directing his/her gaze away from the forward scene and in the first frame of that movement he/she was blinking, the blink was counted as a tenth of a second away. If the driver was always looking forward, then these fields were left null, as that category was not applicable.

Secondary Behaviors
Secondary behaviors were coded as follows:
0 = None
10 = Cellular phone: Conversation, in use (Conversation could include listening, talking, or both while using the cellular phone.)
11 = Cellular phone: Reaching for phone (This classification refers to when the driver reached for the handheld phone in order to speak on that phone. If the driver reached for the phone simply to answer the phone, but then commenced speaking while using a headset, the classification was Other (90).
12 = Cellular phone: Dialing phone
20 = Headset, hands-free phone: Conversation (This was selected when the reviewer could tell that the driver was in a conversation on a hands-free phone.)
21 = Headset, hands-free phone: Reaching for headset
22 = Headset, hands-free phone: Unsure of activity level (The driver was wearing a

headset but it was not clear whether the headset was in use. The driver may have been listening to someone or wearing it in case of an incoming call.)

30 = Eating: High involvement (includes eating a burger, unwrapping food, or other kinds of eating that involve one or both hands off the steering wheel for an extended period of time)

31 = Eating: Low involvement (includes eating candy, grabbing chips, and so forth, where the driver's hands were not necessarily off the steering wheel for an extended period of time)

40 = Drinking: High involvement (includes situations where the driver was trying to open a straw or bottle, blowing on a hot drink, etc. As with eating, the extent to which the driver's hands were off the steering wheel was also a factor)

41 = Drinking: Low involvement (includes situations where the driver was sipping a drink, drinking without looking, etc.)

50 = Conversation (The driver and someone in the car are carrying on a conversation. The driver can be listening during the clip, talking during clip, or doing both.)

60 = In-car system use (The driver was actively adjusting something within the car, usually on or around the front console. For example, the driver was not just listening to the stereo; the driver was also adjusting the stereo, etc. Using the car cigarette lighter was coded under the smoking section.)

70 = Smoking: Lighting (includes the in-car lighter or other means of lighting a cigarette, cigar, etc.)

71 = Smoking: Reaching for cigarettes or lighter or ashtray (includes the in-car lighter)

72 = Smoking (Actively smoking.)

80 = Grooming: High involvement (includes applying makeup, brushing hair, etc. As with eating and drinking, driver hand location was a factor in determining the level of involvement.)

81 = Grooming: Low involvement (includes scratching, running one's fingers through one's hair, etc.)

90 = Other/multiple behaviors, specified in the notes section (These included behaviors that did not fit into any of the other categories, or situations in which the driver was engaged in more than one behavior, all of which were then recorded in the notes section.)

Notes

A notes section recorded any unusual events or ambiguous situations not covered by categories for a particular question. This section also contains general notes on the video clip if there was anything significant taking place that was not adequately covered by the coding process.

APPENDIX B – SECOND UMTRI CODING SCHEME

1	Use Cell Phone	Begin	End
1.1	Prepare to use cell phone	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for phone	Subject initiates another subtask with the cell phone
1.2	Dial phone - Handheld	Subject presses first button	Subject initiates another subtask with the cell phone
1.3	Dial phone – Hands-free	Subject speaks first word	Subject initiates another subtask with the cell phone
1.4	Converse on cell phone (talk, listen)	Subject waits for a response (number is already dialed, phone is at ear)	Subject presses End button or closes phone
1.5	Hold cell phone	Subject holds phone in hand (no activity is taking place with the cell phone)	Subject initiates another subtask with the cell phone
1.6	Hang up cell phone/end call	Subject takes phone from ear or presses End button	Subject puts phone down or initiates another subtask
1.7	Answer cell phone	Subject reaches for phone upon hearing it ring	Subject holds phone in hand and answers call or initiates another subtask

2	Eat/Drink	Begin	End
2.1	Prepare to eat	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for food	Subject initiates another subtask with the food
2.2	Prepare to drink	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for drink	Subject initiates another subtask with the drink
2.3	Eat/bite food - not wrapped	Subject opens mouth	Subject closes mouth
2.4	Eat/bite food - wrapped	Subject opens mouth	Subject closes mouth
2.5	Chew food	Subject moves jaw (to grind food)	Subject swallows food (or subject stops moving jaw)

2.6	Drink from straw or sip from opening (includes can, bottles)	Subject opens mouth	Subject swallows (or subject closes mouth)
2.7	Drink from open top container (cup)	Subject opens mouth	Subject swallows (or subject closes mouth)
2.8	Finish eating	Subject moves to put away wrappers or uneaten food	Subject returns hand to a resting position or initiates another subtask
2.9	Finish drinking	Subject takes cup or container from mouth for the last time (to set it down or dispose of it)	Subject returns hand to a resting position or initiates another subtask
2.10	Spill/drop food	Foods slips from subject's grasp	Subject returns hand to a resting position or initiates another subtask
2.11	Spill/drop drink	Drink slips from subject's grasp	Subject returns hand to a resting position or initiates another subtask
2.12	Hold food/drink	Subject holds food/drink in hand (no other activity is taking place with the food or drink)	Subject returns hand to a resting position or initiates another subtask

3	**Smoke**	**Begin**	**End**
3.1	Prepare to light cigar/cigarette	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for lighter or cigar/cigarette	Subject initiates another subtask with the cigar/cigarette
3.2	Light cigar/cigarette	Subject attempts to light the lighter	Subject pulls lighter away from cigar/cigarette
3.3	Smoke cigar/cigarette	Subject draws on cigar/cigarette	Subject removes cigar/cigarette from mouth for the final time
3.4	Finish smoking	Subject removes cigar/cigarette from mouth for the final time	Subject puts cigar/cigarette out and returns hand to a resting state
3.5	Hold cigar/cigarette	Subject holds cigar/cigarette in hand, or holds in mouth and does not draw on it	Subject initiates another subtask with the cigar/cigarette
3.6	Ash cigar/cigarette	Subject moves hand holding cigar/cigarette to ashtray or window	Subject moves hand to a resting position or initiates another subtask

4	Chew Tobacco	Begin	End
4.1	Prepare to chew tobacco	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for tobacco	Subject places tobacco in mouth
4.2	Chew tobacco	Subject' mouth is closed	Subject moves hand from resting position (steering wheel, lap, etc.) to dispose of tobacco (spittoon, window, etc.)
4.3	Spit (chewing tobacco in mouth)	Subject moves hand from resting position to reach for spittoon, or subject spits (through open window)	Subject returns hand to a resting position
4.4	Remove chewing tobacco from mouth	Subject moves hand from a resting position to remove the tobacco from mouth	Subject moves hand to a resting position or initiates another subtask

5	Chew Gum	Begin	End
5.1	Hold gum in mouth	Subject's mouth is static	Subject initiates another subtask with the gum
5.2	Prepare to chew gum	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for gum	Subject places piece of unwrapped gum in mouth
5.3	Blow gum bubble	Subject stretches gum	Bubble pops
5.4	Remove popped gum bubble	Subject moves to collect gum	Subject has all gum in mouth
5.5	Chew gum	Subject lowers jaw	Subject's jaw is static
5.6	Bite/lick lips	Subject moves lips/tongue	Subject's lips/tongue are at rest
5.7	Tongue motion	Subject moves tongue (excludes tongue motion to keep gum in place)	Subject's tongue returns to a resting state or subject closes mouth (tongue inside mouth)
5.8	Finish chewing gum	Subject moves to take gum from mouth or spit gum out	Subject returns head/hand to a resting position
5.10	Other - chewing gum	Subject begins other gum related activity	Subject ends other gum related activity

6	**Groom**	**Begin**	**End**
6.1	Prepare to groom	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for grooming tool or to perform grooming task with hand	Subject initiates another grooming subtask
6.2	Groom - hand only	Subject touches grooming area with hand	Subject removes hand from grooming area
6.3	Groom - using tool	Subject touches grooming area with grooming tool	Subject removes hand holding grooming tool from grooming area
6.4	Hold grooming tool	Subject holds grooming tool in hand while not touching the grooming area	Subject initiates another grooming subtask
6.5	Finish grooming	Subject removes hand or grooming tool from grooming area	Subject moves hand to a resting position or initiates another subtask

Note: Items include Notepad, Newspaper, Magazine, Maps, etc.

7	**Read**	**Begin**	**End**
7.1	Prepare to read	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for reading material	Subject initiates another reading subtask
7.2	Read	Subject opens reading material	Subject initiates another reading subtask
7.3	Put away/fold reading materials	Subject moves to close reading material	Subject moves hand to a resting position or initiates another subtask

Note: Items include Notepad, Newspaper, Magazine, Maps, etc.

8	**Write**	**Begin**	**End**
8.1	Prepare to write	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for writing utensil	Subject initiates another writing subtask
8.2	Write	Subject touches writing utensil to writing surface	Subject initiates another writing subtask
8.3	Put away writing materials	Subject moves to put away writing utensils	Subject moves hand to a resting position or initiates another subtask

9	Type	Begin	End
9.1	Prepare to type	Subject moves hand from resting position (steering wheel, lap, etc.) to reach for device	Subject initiates another typing subtask
9.2	Type with 1 thumb	Subject types first character	Subject initiates another typing subtask
9.3	Type with 2 thumbs	Subject types first character	Subject initiates another typing subtask
9.4	Type on full keyboard	Subject types first character	Subject initiates another typing subtask
9.5	End typing	Subject types last character	Subject returns hand to a resting position or initiates another subtask

10	Use In-Car System	Begin	End
10.1	No adjustment of in-car system		
10.2	Control steering wheel	Subject moves hand from resting position to turn steering wheel	Subject returns hand to a resting position or initiates another subtask
10.3	Control stalk	Subject moves hand from resting position (steering wheel, lap, etc.) to control stalk	Subject returns hand to a resting position or initiates another subtask
10.4	Control IP, column, or center console	Subject moves hand from resting position (steering wheel, lap, etc.) to control IP, column or center console	Subject returns hand to a resting position or initiates another subtask
10.5	Control door	Subject moves hand from resting position (steering wheel, lap, etc.) to control door	Subject returns hand to a resting position or initiates another subtask
10.6	Glance only - monitor in-car system	Subject glances away from road	Subject returns attention to the road
10.7	Other or unknown - in-car system use	Subject moves hand from resting position (steering wheel, lap, etc.), to control unknown device	Subject returns hand to a resting position or initiates another subtask

11	Internal Distraction	Begin	End
11.1	Catch falling object/prevent object from moving, reach/lean/look for/pick up	Subject moves hand from resting position (steering wheel, lap, etc.), to reach for object	Subject returns hand to a resting position or initiates another subtask
11.2	Insect-related distraction	Subject moves hand from resting position (steering wheel, lap, etc.), to attend to insect	Subject returns hand to a resting position or initiates another subtask
11.3	Pet related distraction	Subject moves hand from resting position (steering wheel, lap, etc.), to attend to pet	Subject returns hand to a resting position or initiates another subtask
11.4	Glance only - monitor internal distraction	Subject glances away from road	Subject returns attention to the road
11.5	Other - internal distraction	Subject moves hand from resting position (steering wheel, lap, etc.), to attend unknown internal distraction	Subject returns hand to a resting position or initiates another subtask

12	Drowsy	Begin	End
12.1	Close eyes slowly - drowsy	Subject's eye/eyes begin to close slowly (not a blink)	Eye/eyes return to fully opened state
12.2	Head dip - drowsy	Subject's head begins to lower involuntarily	Subject returns head to an upright position
12.3	Yawn	Mouth begins to open to yawn	Subject closes mouth fully

13	Converse	State
13.1	Converse with unknown	Subject converses, but subject's eyes or head is not focused toward a discernable passenger
13.2	Converse with passenger - speak	Subject speaks to a passenger
13.3	Converse with passenger - listen	Subject listens to a passenger speak (passenger is talking to the driver
13.4	Sing/talk to self	Subject sings/talks to himself/herself. There is no passenger in the car and subject is not using a cell phone
13.5	Talk to someone outside vehicle (not by phone)	Subject yells/converses with person outside vehicle through the driver's side window
13.6	Road rage	Subject is visibly agitated (may be talking to self or passenger, may be yelling)

14	**Eyes**
14.1	Looking forward at forward scene
14.2	Looking at left outside mirror or left window
14.3	Looking back over left shoulder
14.4	Looking at right outside mirror or right window
14.5	Looking back over right shoulder
14.6	Looking forward at rear-view mirror
14.7	Eyes down, looking at instrument panel or at lap area
14.8	Eyes down, looking at center stack counsel area
14.9	Transition (eyes not focused on anything)
14.10	Cannot evaluate eye location (sunglasses, glare, etc.)
14.11	Blink (eyes closed)
14.12	Other - Eyes

15	**Head**
15.1	Head facing forward at forward scene
15.2	Head turned toward left outside mirror or left window
15.3	Head turned back over left shoulder
15.4	Head turned toward right outside mirror or right window
15.5	Head turned back over right shoulder
15.6	Head aimed forward toward rear-view mirror
15.7	Head down, turned toward instrument panel
15.8	Head down, turned toward center stack counsel area
15.9	Head down, turned toward lap area
15.10	Transition (head in motion, between positions)
15.11	Other - Head

16	**Hands**
16.1	Cannot evaluate hand position
16.2	Both hands on steering wheel
16.3	1 hand on steering wheel, other unknown
16.4	1 hand performing action, other unknown
16.5	1 hand performing action, other on steering wheel
16.6	1 hand performing 2 or more actions, other unknown
16.7	1 hand performing 2 or more actions, other on steering wheel
16.8	2 hands performing 1 action (not on steering wheel)
16.9	2 hands performing 2 or more actions (not on steering wheel)

16.10	2 hands on wheel, 1 performing action
16.11	2 hands on wheel, both performing action
16.12	2 hands on wheel, performing 2 or more actions
16.13	One hand idle/resting, other unknown
16.14	One hand idle/resting, other on wheel
16.15	One hand supporting head, other unknown
16.16	One hand supporting head, other on wheel
16.17	One hand on wheel AND performing action, other unknown

Driving Context Codes Used in This Study

17 **Weather/Visibility**

17.1	Clear
17.2	Raining
17.3	Snowing/Sleeting

18 **Surface Conditions**

18.1	Dry
18.2	Wet
18.3	Snowy

Driving Context Codes for Future Studies

17 **Weather/Visibility**

17.1	Clear
17.2	Cloudy
17.3	Fog
17.4	Mist
17.5	Raining
17.6	Snowing
17.7	Sleeting
17.8	Smoke/Dust
17.9	Other

18 **Surface Conditions**

18.1	Dry
18.2	Wet
18.3	Snowy
18.4	Icy
18.5	Muddy
18.6	Oily
18.7	Other

APPENDIX C – DISTRACTION FREQUENCIES FROM INITIAL ACAS DATA ANALYSIS

Findings From the ACAS Study, Week 1 (no ACC or FCW)
(Ervin, Sayer, LeBlanc, Bogard, Mefford, Hagan, Bareket, Winkler, 2005)

Non-driving Behavior	Female Younger	Female Middle-aged	Female Older	Male Younger	Male Middle-aged	Male Older	Total Clips	% of Total Week 1 Clips
Cell phone: conversation, in use	4			5	2		11	7
Cell phone: reaching for							0	0
Cell Phone: dialing							0	0
Conversation		3	1		2	1	7	4
Drinking: high involvement							0	0
Drinking: low involvement					1		1	1
Eating: high involvement							0	0
Eating: low involvement		1					1	1
Grooming: high involvement							0	0
Grooming: low involvement		1		1	1	2	5	3
Headset/hands-free phone: conversation							0	0
Headset/hands-free phone: reaching for headset							0	0
Headset/hands-free phone: unsure if any activity				2			2	1
In-car system use							0	0
None	22	25	22	30	14	19	132	83
Other/multiple behaviors							0	0
Smoking: lighting a cigarette							0	0
Smoking: reaching for cigarettes or lighter							0	0
Smoking						1	1	1
Total Clips	26	30	23	38	20	23	160	
Clips w/ non-driving behaviors (%)	4 (15)	5 (17)	1 (4)	8 (21)	6 (30)	4 (17)	28 (18)	

APPENDIX D – CODING GUIDELINES

1. How to Code:

- Since there is no audio available, conversation could only be determined using mouth movement, and it is very important not to confuse this with various forms of chewing.

- Defining start and end of conversation is complex in the absence of audio. For the purposes of the SAVE-IT project, if mouth movement is subtle, conversation only starts with the opening of the mouth and ends with the final open-mouthed movement. However, if the initial mouth movement is obvious, the last static frame before the movement counts toward speech. If at the end of speech, the last frame shows that the mouth undergoes significant closure, the first static frame also counts as speech. This criteria is somewhat arbitrary, and is used to balance the fact that audio could not be used to distinguish when a speaker vocalizes a lipped consonant (the letters b, m, and p). Observation of facial muscles is crucial to aid this judgment. If audio were present, conversation could simply be defined as the period of time the speaker is audibly forming words or sounds (e.g., um, uh).

- If a behavior was noticed in a clip, every instance of the behavior was recorded only in those frames in which it was noticed, not continuously throughout the clip. There are exceptions to this principle. For example, if the hand was out of view, but the cup it was holding was still seen, the hand was assumed to be holding the cup. Also, if a behavior disappeared from the camera view for only one or two frames, context could also support the assumption that the behavior was ongoing.

- It was hard to distinguish between singing, road rage, and conversing with passengers, since the content of the speech could not be discerned. Therefore, “conversation with unknown” is prominent.

- Conversing with unknown/passengers is based on observable speech. Such interaction with a passenger is only supported if:
    - The driver’s speech is synchronized with the visible passenger (driver stops talking right when the passenger begins, and vice versa).
    - The driver glances at the passenger while talking or listening (rear view mirrors do not qualify, as glance directions in such cases are not conclusive).
    - The driver or passenger physically leans, as if to listen better or acknowledge the speaker.
    - If the driver looks to the right side while speaking, conversation is not coded as being with a passenger, even if a passenger is present. An exception to this is made if there are no distinguishing features on the right side of the road and the driver glances to the right for more than two frames.

- Glance behavior was used to distinguish between internal distraction and in-car system use, as hand position was extremely hard to determine. Shoulder movements were often used to assist hand position assessments.

- Sunglasses were very useful in spotting otherwise covert behavior, such as reading.

- Blinking and transition are difficult to code when the eyes blink and change to another direction. Blinking is defined as the covering of the pupils. As long as this condition was met, blinking always superseded transition.

- Transition:
    - Is defined as head or eye positions between the original static position and the destination static position. A static position is defined as a position where the head or eyes appear to be fixed on an object.
    - Should not be more than 3 frames.
    - The driver is scanning if the eyes and head consistently look in similar directions, as if they are fixed on a moving object (similar to gradually changing static positions). Scanning behavior is not considered transitioning, and is not coded.
    - Frames where blurring of the eyes or head occur (due to rapid movement of the driver) are counted as transition frames.

- The directions of the eyes and head are to be coded independently.

- Eye closure (blinking) can complicate instances of drowsiness (slow eye closure). Slow eye closure was only recorded as it was seen, even though the person is likely to be constantly drowsy. Blinking is not classified as drowsiness.

- Reflective asphalt can easily be confused with a wet road surface, especially at night. In this case, use other cues to determine if a road is wet; for example, tracks in the road may indicate where the tires have pushed water aside, leaving the surface partially dry. Other cues depend on the time of day and the quality of the video used.

- While snow was relatively obvious on the forward camera, rain was much harder to spot. Active windshield wipers, streaming water on the rear window, and excessive glare were all used as cues for rain, as was fresh rainfall on the windshield.

2. Comments and Suggestions:

- Improvements to the data collection equipment are needed. A frame rate of 10 Hz to show a smooth frame rate is desirable. Also, a finer resolution and color would help the coder identify rain and the condition of the road surface.

- Additional camera views would help identify hand position, in-car system use, internal distraction, reading, writing, and typing. Suggested additional views include a camera to view the area around the instrument panel, steering wheel, and center column stack. Another camera placed right next to the right A-pillar would help the coder to classify whether the driver is glancing out the right mirror or the forward view; this would also help decide if the driver is trying to look over his or her shoulder.

- Certain glance angles were difficult to code with the current system. A slight change could transform a glance at the forward scene to a glance at the center mirror. Also, the change from a forward glance to a glance out the right mirror could be tough to pinpoint if the driver is scanning.