A TAXONOMY FOR SECONDARY IN-VEHICLE TASKS BASED ON EYE GLANCE AND TASK COMPLETION TIME

Myra Blanco and Jonathan M. Hankey
Virginia Tech Transportation Institute
3500 Transportation Research Plaza (0536)
Blacksburg, VA 24061

Jacqueline A. Chestnut
General Motors Corporation
30200 Mound Road
Warren, MI 48090-9010

ABSTRACT

The objective of this research was to develop an initial taxonomy that grouped similar secondary in-vehicle tasks based on driving-related performance measures. This type of taxonomy would be useful to system designers when developing in-vehicle tasks and to researchers. Research was conducted using 2 infotainment systems, 17 tasks, and 89 participants to develop and validate an initial taxonomy. The results indicate that the 17 tasks could be parsed into four distinct groups ranging from selecting an AM band to destination entry. The groupings are based on number of glances and task completion time, which provided the best separation between the groups and consistent results for both static and dynamic testing.

INTRODUCTION

Background

Research and evaluation on secondary in-vehicle tasks have been performed for decades (Wierwille, Gutmann, Hicks, & Muto, 1977). However, in the last decade secondary in-vehicle tasks and their complexity have increased (Blanco, Biever, Gallagher, & Dingus, submitted; Tsimhoni, Smith, & Green, 2004). Modeling, testing and evaluation, and research efforts are often performed with a subset of data that might not represent a given system's full spectrum of tasks. As new in-vehicle systems are introduced, it is important for designers, engineers, and researchers to evaluate the full spectrum of tasks to avoid misrepresenting how drivers will behave under several levels of task complexity.

Taxonomies have been developed to help researchers and practitioners better understand various issues such as pedestrian injuries (Schofer et al., 1995), safety-related work behavior (DeJoy, 1993), warnings (Ayres, Gross, Horst, & Robinson, 1992), mental models (Moray, 1996), pilot training tasks (Meyer, Laveson, Weissman, & Eddowes, 1975), and web design (Cheong & Shehab, 2003) to name a few. However, the development of detailed taxonomies for driver-related issues is not as abundant. The existing driver-related taxonomies deal with issues such as driver distraction (Donmez, Boyle, & Lee, 2003; Mackie & Wylie, 1991), driver error (Hankey et al., 1999), and older drivers (Hanowski, Bittner, Knipling, Byrne, & Parasuraman, 1995) but not with secondary in-vehicle tasks. Consequently, a standardized taxonomy to identify the different layers of complexity of secondary in-vehicle tasks to be used for research and evaluation purposes is a gap in the transportation human factors field that should be explored.

Several characteristics need to be considered to develop a useful taxonomy identifying the complexity of different secondary in-vehicle tasks: sensitivity, level of intrusion, diagnosticity, transferability, and implementation requirements (Wierwille and Eggemeier, 1993). The sensitivity of the taxonomy should be sufficient to clearly discriminate between different levels of task complexity without causing changes in how the driver interacts with the system of interest (i.e., non-intrusive). The taxonomy should help diagnose the level of information processing needed to perform the task. It should be easy to use independently of the system being evaluated (i.e., transferable). The taxonomy needs to be capable of providing a means of evaluation early in system development when simulator or on-road testing is not an option. Therefore, these measurements should be capable of predicting task complexity in a dynamic environment (i.e., driving) as well as in a static environment (e.g., early stage of a prototype such as bench testing).

A taxonomy that identifies levels of secondary in-vehicle tasks complexity should serve as a link that connects these different layers of tasks to potential variations in driving performance. Several measurements have been identified in the literature as being sensitive to changes in drivers' workload. For example, speed and accuracy of performance (e.g., lane maintenance) are expected to be more erratic as workload increases (Wierwille and Eggemeier, 1993). Because the primary goal is to identify secondary tasks that affect the primary task of driving, the taxonomy should be strongly based in measurements that can accurately predict changes in driving performance. In order to assess driving safety, measurements that correlate secondary task complexity to driving performance should be used. The most commonly cited measurements that require minimal implementation are: eye glance frequency, task completion time, and number of steps to complete the task (Alliance of Automotive Manufacturers, 2003; Angell, Young, Hankey & Dingus, 2002; Green, 1995; Society of Automotive Engineers, 2002, 2004).

Research Objectives

The objective of this research effort was to find an empirical taxonomy that is easily applicable to the full gamut of secondary in-vehicle tasks and uses metrics that produce meaningful and comparable results in both static and dynamic environments.

METHODS

Participants

A total of 89 drivers between the ages of 25 and 65 years old participated in this research effort. The study used two different instrumented vehicles (with different infotainment systems), one for the taxonomy development portion and a second for validation purposes. Participants were randomly assigned to one of the two vehicles, controlled by age and gender for both. Participants were also divided into two conditions: Static and Dynamic.

Apparatus

Two static stations were set up within one of the garages of the Virginia Tech Transportation Institute (VTTI) facility. Each station enclosed a vehicle using black curtains. Within each static station, a 27-in. color television was placed on a cart in front of the vehicle and connected to a VCR. A looped, pre-recorded video of a driving route was shown throughout the static testing to enhance the reality of the static scenario. The vehicles were left running, and their parking lights were left on during the static testing. Exhaust fumes were vented outside the building.

The dynamic portion of the study took place on the Smart Road testing facility at VTTI. This testing facility is a two-lane roadway approximately 2.1 miles long. In the dynamic portion participants drove the Smart Road while performing secondary tasks.

The vehicles were instrumented for continuous data collection. For static and dynamic apparatus, the data-collection system consisted of video cameras to record pertinent environmental events and eye-movement data. In the dynamic portion, these variables were collected as well as speed and lane deviation information. The sampling rate was 10 Hz for the vehicle performance measures and 30 Hz for video recording. The data stream was flagged by the experimenter when a task occurred for post processing and analysis purposes.

Experimental Design

The design used for the experiment was a within-participant design with one independent variable: task type. Seventeen different types of secondary in-vehicle tasks were performed. The task presentation order was counterbalanced, and one task trial was performed for each task:

1. **Radio/CD:** Change radio band to AM from CD
2. **Radio/CD:** Tune radio to 710 AM
3. **Radio/CD:** Go to FM2 preset radio station 5
4. **Radio/CD:** Set current CD to track 26
5. **Radio/CD:** Select next track on current CD
6. **HVAC:** Adjust the fan speed up two levels
7. **HVAC:** Change temperature from 72 °F to 70 °F
8. **Cell Phone:** Dial 10-digit home phone number on hand-held cell phone
9. **Settings:** Adjust display daytime/nighttime setting to AUTO
10. **Edit Route:** Edit the calculated route to E. Main St., Martinsville, VA to minimize tolls
11. **Map:** Change map appearance to Turn List (destination already available)
12. **Map:** Change navigation area from Michigan (MI) to Virginia (VA)
13. **Map:** Zoom in map 5 units
14. **Destination List:** Get directions to 1100 N. Main St, Ann Arbor, MI from the previous destination list
15. **Full Address:** Get directions to 2375 Brickley in Ferndale, MI
16. **Intersection:** Get directions to the intersection of Elmwood Avenue and Oakwood Street in Ann Arbor, MI
17. **Point of Interest:** Get directions to Taco Grande Mexican Restaurant in Kokomo, IN

Tasks that represent what is currently available in the market were selected (including advanced in-vehicle information systems). However, the length of the experimental session and the driver's potential fatigue are important aspects that need to be taken into consideration when designing a comprehensive study. Therefore, a single repetition was deemed appropriate for this study; this was compensated for by using a larger sample to maintain statistical power. In addition, thorough training was given to the participants on the systems of interest prior to testing.

Procedure

Each participant performed in only one test session. The participant was taken to the test vehicle, oriented to the vehicle controls, and given an overview of the infotainment system. Next, orientation tasks, presented in the same order for all participants, were administered to familiarize participants with the testing procedures.

Each participant then performed all the tasks in a counterbalanced order. Prior to performing each task, participants were trained on the specific task so it could be performed without assistance during data collection. This procedure—training, task performance, and training for the next task—was followed until data had been collected on all the tasks.

Dependent Variables

Three different types of dependent measurements were obtained: (1) eye glance, (2) vehicle control (dynamic only), and (3) secondary task performance.

In terms of eye glance behavior, the number of glances to the system of interest (i.e., infotainment, cell phone, HVAC) were collected:

- Number of Glances: A glance was defined as all consecutive fixations to the system of interest and any preceding transition. For example, during a radio-related task, all fixations and saccades towards the radio interface at the center stack before a transition to another location was defined as one glance to the radio. Glance data were recorded in real-time. Video data were post-processed, categorizing glance locations during each task.

For vehicle control, several measurements were collected:

- Number of lane deviations: Defined as when the vehicle's tire came into contact with the lane marker.
- Time out of the lane: Defined as the time from when the tire came into contact with the lane marker until the tire was no longer in contact with the lane marker and the vehicle was in the correct lane.
- Percent of time out of lane: The percentage of time the participant was out of the correct driving lane.
- Speed variance: The variance in vehicle speed during a task.

The measurements of interest for secondary task performance were:

- Number of steps to perform the task: A task step is a measurable singular action or procedure that has an observable beginning and end.
- Task completion time: Measured from the time the experimenter said "Begin" until the participant said "Done."

RESULTS

In order to determine if a dependent variable measure produced a naturally emerging taxonomy, several conditions needed to be fulfilled. First, the measurement needed to be both significant and have a high correlation ($r > 0.70$) between static and dynamic conditions. Second, a one-way ANOVA ($\alpha = 0.05$) needed to show that the measurement was able to discriminate between the different task types (static and dynamic environment data were analyzed separately). A Student-Newman-Keuls (SNK) post-hoc analysis was performed for the significant main effect ($p < 0.05$) to reveal the natural groupings. Third, significant measurements were compared on their classification patterns based on the SNK.

The correlation analysis showed that measurements like number of glances, number of steps, and task completion time have strong correlations between static and dynamic environments (Table 1). The ANOVA results suggest that all the measurements, with the exception of percent of time out of lane, can discriminate between the tasks (all significant measurements had a $p < 0.001$). The SNK suggested natural groupings, creating a taxonomy of secondary in-vehicle tasks for both static and dynamic environments. Based on the conditions described, the two measurements that clearly defined the taxonomy were: (1) number of eye glances to the system of interest and (2) task completion time (Table 2). The suggested taxonomy (Table 2) was validated with the results obtained from the second vehicle.

Number of steps had a significant ($p < 0.05$) and high-strength ($r > 0.70$) correlation between the static and dynamic environments. However, SNK was not able to differentiate between the different types of tasks. For example, even though the number of analytical steps required to complete a dialing task was less than half of the number of steps needed for getting to track 26 on a CD (i.e., 12 and 26 steps, respectively), the number of eye glances were, on average, fairly similar. This might be due to the fact that dialing the phone requires the driver to perform keystrokes to different, precise locations, while getting to track 26 on a CD requires the use of just one key to advance to the track of interest as well as occasional checking of the current track number. This is confirmed by the fact that the difference in terms of number of glances to the system of interest between going to the next track on a CD (i.e., 2.0 glances) and going to track 26 (i.e., 6.2 glances) is not a one-to-one multiplier.

Table 1. Correlation of static and dynamic measurements.

		Static					
		No. glances		No. steps		Task time	
Dynamic	No. glances	0.959	*	0.927	*	0.954	*
	No. steps	0.979	*	0.953	*	0.973	*
	Task time	0.950	*	0.916	*	0.945	*
	No. lane deviations	0.975	*	0.957	*	0.965	*
	Speed variance	0.926	*	0.882	*	0.918	*
	Percent of time out of lane	0.468		0.461		0.444	
	Time out of lane	0.970	*	0.942	*	0.955	*

* = $p < 0.05$ (significant)

It is important to understand that a task step involves more than a keystroke or a knob turn. Depending on the user's goal, information gathering and complex decision-making might be required prior to the next step. Tasks such as the navigation task in Group 3 (i.e., changing the navigation area) and the tasks in Group 4 (i.e., entering directions) require a significant amount of cognitive processing even though they require fewer analytical task steps to be completed than getting to track 26 on a CD. These tasks entail combinations of several complex decision-making elements, such as searching, identifying, and interpreting (Lee, Morgan, Wheeler, Hulse, & Dingus, 1997), using multiple screens, and viewing high information density in each screen. Not surprisingly, complex decision-making and high information density have shown to often increase visual attention demand for in-vehicle information-system-related tasks (Blanco, 1999; Blanco, Biever, Gallagher, & Dingus, submitted; Gallagher, 2001).

Therefore, number of steps is not suggested as part of this taxonomy. Based on data comparisons similar to the one discussed above for the CD and cell phone tasks, number of steps was not deemed an accurate descriptor of the complexity of a task at least in terms of attention demand and cognitive complexity.

Table 2. Taxonomy for secondary in-vehicle tasks based on number of glances and task completion time.

	Glances	Time (sec)	Type of Task [Analytical Steps]
Group 1	g < 3	t < 7	• **Radio/CD:** Change radio band to AM from CD **[1]** • **Radio/CD:** Select next track on current CD **[2]** • **HVAC:** Adjust the fan speed up two levels **[3]** • **HVAC:** Change temperature from 72 °F to 70 °F **[2]** • **Map:** Change map appearance to Turn List (destination already available) **[1]**
Group 2	3 < g < 6	7 < t < 15	• **Radio/CD:** Tune radio to 710 AM **[2]** • **Map:** Zoom in map 5 units **[5]** • **Radio/CD:** Go to FM2 preset radio station 5 **[3]** • **Edit Route:** Edit the calculated route to E. Main St., Martinsville, VA to minimize tolls **[2]** • **Settings:** Adjust display daytime/nighttime setting to AUTO **[3]** • **Destination List:** Get directions to 1100 N. Main St, Ann Arbor, MI from the Previous Destination list **[4]**
Group 3	6 < g < 9	15 < t < 25	• **Cell Phone:** Dial 10-digit home phone number on hand-held cell phone **[12]** • **Radio/CD:** Set current CD to track 26 **[26]** • **Map:** Change navigation area from Michigan (MI) to Virginia (VA) **[6]**

Group 4	g > 9	T > 25	• **Destination Entry**
			  ○ full address **[15]**
			  ○ intersection of two streets **[12]**
			  ○ point of interest **[17]**

DISCUSSION AND CONCLUSIONS

The results indicate that a taxonomy based on number of eye glances to the system of interest together with task completion time shows the most promise. Both of these measurements are accepted in the transportation safety arena as important metrics and have shown correlation to driving performance in this study as well as in past research (Angell, Young, Hankey & Dingus, 2002; Green, 1999a, 1999b; Tijerina, Palmer, & Goodman, 1998). Furthermore, they are metrics that are easy to apply with minimal processing or data reduction.

The current research also demonstrates reciprocity between static and dynamic environments for this taxonomy, allowing it to serve as a practical tool for both designers and researchers. Ultimately it is hope that this taxonomy will allow transportation safety professionals to cover the full spectrum of secondary in-vehicle tasks available. In addition, using a standard taxonomy will improve the communication of results between different organizations as well as the external validity of experimental results.

Although the tasks in this taxonomy provide a representative sample of secondary tasks types possible in the driving environment, designers may identify tasks that do not clearly fit into these grouping. In an effort to facilitate task categorization, the general characteristics of the different task groups are described in Table 3 using characteristics discussed in terms of visual, manual, and cognitive demands (Blanco, 1999; Blanco, et al., submitted; Gallagher, 2001; Lee et al., 1997). These characteristics attempt to provide the designer with basic attributes of tasks that could be used to provide an initial categorization. The required level of monitoring, complex decision-making, and long-term memory, as well as the way the information is displayed and accessed all impact the task grouping and difficulty.

Finally, all the tasks evaluated in this research are manual-visual tasks. Future research should expand on this taxonomy to include tasks that are mainly auditory and performed with voice recognition.

Table 3. Characteristics of the different groups of secondary in-vehicle tasks.

	Glances	Time (sec)	General Characteristics
Group 1	g < 3	t < 7	• One display screen • Low information density • Working memory • Decision-making elements used: search, detect, and control with minor monitoring
Group 2	3 < g < 6	7 < t < 15	• One or two display screens • Low information density • Working memory • Decision-making elements used: search, detect, identify, select, and control with minor monitoring
Group 3	6 < g < 9	15 < t < 25	• One or two display screens • Medium information density • Working memory and long-term memory • Decision-making elements used: search, detect, identify, select, and control with an intermediate level of monitoring
Group 4	g > 9	t > 25	• More than two display screens • High information density • Working memory and long-term memory • Decision-making elements used: search, detect, identify, interpret, code, select, and control with major monitoring

REFERENCES

Alliance of Automobile Manufacturers (2003). Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced In-Vehicle Information and Communication Systems (Version 3). Washington, D.C.: Alliance of Automobile Manufacturers.

Angell, L.S., Young, R.A., Hankey, J.M., & Dingus, T.A. (2002). An evaluation of alternative methods for assessing driver workload in the early development of in-vehicle information (2002-01-1981). In *Proceedings of the Government/Industry Meeting*. Warrendale, PA: Society of Automotive Engineers.

Ayres, T. J., Gross, M. M., Horst, D. P., & Robinson, J. N. (1992). A methodological taxonomy for warnings research. In *Proceedings of the Human Factors and Ergonomics Society 36th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Blanco, M. (1999). Effects of In-Vehicle Information Systems (IVIS) Tasks on the Information Processing Demands of a Commercial Vehicle Operations (CVO) Driver. Unpublished masters thesis. Virginia Polytechnic Institute and State University, Blacksburg, VA. URL: http://scholar.lib.vt.edu/theses/available/etd-122299-195616/

Blanco, M., Biever, W. J., Gallagher, J. P., & Dingus, T. A. (2005). *The impact of secondary task cognitive processing demand on driving performance.* Manuscript submitted for publication.

Cheong, Y., & Shehab, R. L. (2003). Building a web-features taxonomy for structuring web design guidelines. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

DeJoy, D. M. (1993). Development of a work behavior taxonomy for the safety function in industry. *Accident Analysis & Prevention, 25,* 365.

Donmez, B., Boyle, L., & Lee, J. D. (2003). Taxonomy of mitigation strategies for driver distraction. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Gallagher, J. P. (2001). An assessment of the attention demand associated with the processing of information for in-vehicle information systems (IVIS). Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA. URL: http://scholar.lib.vt.edu/theses/available/etd-05032001-092804/

Green, P. (1995). Measures and Methods Used to Assess the Safety and Usability of Driver Information Systems (Technical Report UMTRI-93-12). Ann Arbor, MI: University of Michigan Transportation Research Institute.

Green, P. (1999a). Visual and Task Demands of Driver Information Systems (Technical Report UMTRI-98-16). Ann Arbor, MI: University of Michigan Transportation Research Institute.

Green, P. (1999b). The 15-second rule for driver information systems. In *Proceedings of the ITS America 9th Annual Meeting*. Washington, DC: ITS America.

Hankey, J. M., Wierwille, W. W., Cannell, W. J., Kieliszewski, C. A., Medina, A., Dingus, T. A., & Cooper, L. M. (1999). *Identification and evaluation of driver errors: Task C report, driver error taxonomy development* (Report under Contract No. DTFH61-97-C-00051). Blacksburg, VA: Virginia Tech Transportation Institute.

Hanowski, R. J., Bittner, A. C., Knipling, R. R., Byrne, E. A., & Parasuraman, R. (1995). Analysis of older driver safety interventions: A human factors taxonomic approach. In *Proceedings of the Intelligent Transportation Society of America 3rd Annual Meeting and Exposition*. Washington, DC: Intelligent Transportation Society of America.

Lee, J.D., Morgan, J., Wheeler, W.A., Hulse, M.C., & Dingus, T.A. (1997). Development of Human Factors Guidelines for Advanced Traveler Information Systems (ATIS) and Commercial Vehicle Operations (CVO): Description of ATIS/CVO Functions (FHWA-RD-95-201). Washington D.C.: Federal Highway Administration.

Mackie, R. R., & Wylie, C. D. (1991). Countermeasures to loss of alertness in motor vehicle drivers: A taxonomy and evaluation. In *Proceedings of the Human Factors and Ergonomics Society 35th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Meyer, R. P., Laveson, J. I., Weissman, N. S., & Eddowes, E. E. (1975). Behavioral taxonomy of undergraduate pilot training tasks and skills. In *Proceedings of the Human Factors and Ergonomics Society 19th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Moray, N. (1996). A taxonomy and theory of mental models. In *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Schofer, J. L., Christoffel, K. K., Donovan, M., Lavigne, J. V., Tanz, R. R., & Wills, K. E. (1995). Child pedestrian injury taxonomy based on visibility and action. *Accident Analysis & Prevention, 27,* 317.

Society of Automotive Engineers (2002). SAE Recommended Practice: Calculation of the Time to Complete In-Vehicle Navigation and Route Guidance Tasks (SAE J2365). Warrendale, PA: Society of Automotive Engineers.

Society of Automotive Engineers (2004). SAE Recommended Practice Navigation and Route Guidance Function Accessibility While Driving (SAE 2364). Warrendale, PA: Society of Automotive Engineers.

Tijerina, L., Palmer, E., & Goodman, M.J. (1998). Driver workload assessment of route guidance system destination entry while driving: A test track study. In *Proceedings of the 5th ITS World Congress*. Washington, DC: ITS America.

Tsimhoni, O., Smith, D., & Green, P. (2004). Address entry while driving: Speech Recognition versus a touch-screen keyboard. *Human Factors,46,* 600.

Wierwille, W. W., & Eggemeier, F.T. (1993). Recommendations for mental workload measurements in a test and evaluation environment. *Human Factors, 35,* 263.

Wierwille, W. W., Gutmann, J. C., Hicks, T. G., & Muto, W. H. (1977). Secondary task measurements of workload as a function of simulated vehicle dynamics and driving conditions. *Human Factors, 19,* 557.