COGNITION and NEUROERGONOMICS (CaN)

COLLABORATIVE TECHNOLOGY ALLIANCE (CTA)

ARMY RESEARCH LABORATORY (ARL)

STANDARDIZED ANNOTATED NEUROPHYSIOLOGICAL DATA REPOSITORY (SANDR)

DATA DESCRIPTION v2.2.2

PREPARED BY:

DCS Corporation

PREPARED FOR:

Combat Capabilities Development Command
Army Research Laboratory
Human Research and Engineering Directorate
Building 459 Mulberry Point Road
Aberdeen Proving Ground, MD 21005

This page intentionally left blank

Table of Contents

1 Data Repository Overview ... 1
1.1 SANDR Summary Document ... 1
1.1.1 Revision History ... 1
1.1.2 Acknowledgements ... 1
1.2 Data Repository Description ... 2
1.2.1 Data Standardization Process ... 3
1.3 SANDR-related Publications ... 20
1.4 Summary of Datasets ... 21
1.4.1 Repeat subjects ... 25
2 Study Descriptions ... 27
2.1 Brain-Computer Interface Technology (BCIT) ... 27
2.1.1 BCIT Program Summary ... 28
2.1.2 BCIT Publications ... 46
2.1.3 BCIT Datasets ... 48
2.2 High Definition Cognition (HDCOG) ... 51
2.2.1 HDCOG Program Summary ... 52
2.2.2 HDCOG Publications ... 68
2.2.3 HDCOG Datasets ... 68
2.3 Institute for Collaborative Biotechnologies (ICB) ... 69
2.3.1 ICB Program Summary ... 70
2.3.2 ICB Publications ... 78
2.3.3 ICB Datasets ... 79
2.4 Cognition and Neuroergonomics Collaborative Technology Alliance (CaN CTA) ... 80
2.4.1 CaN CTA Program Summary ... 80
2.4.2 CANCTA Publications ... 128
2.4.3 CANCTA Datasets ... 133
2.5 Autonomy Research Pilot Initiative (ARPI) ... 134
2.5.1 ARPI Program Summary ... 136

2.5.2 ARPI Publications 159
2.5.3 ARPI Datasets 159

List of Figures

Figure 1-1. Traditional EEG Analysis Path versus Big Data Analysis 2
Figure 1-2. Big Data for Human Sensing EEG Pipeline Process 3
Figure 1-3. Big Data for Human Sensing System Architecture 5
Figure 1-4. Sample Raw Data Folder (TX16) 6
Figure 1-5. TX16 Data Collection System Architecture 6
Figure 1-6. BioSemi Channel Location Remap 7
Figure 1-7. EEG to SIM Sync Plot (TX16) 8
Figure 1-8. Task Paradigm and Tags 9
Figure 1-9. Hierarchical Event Descriptor 10
Figure 1-10. Sample HED Tags 11
Figure 1-11. Sample HED Tag 12
Figure 1-12. HED Validator 13
Figure 1-13. EEG.event Structure with HED Tags 14
Figure 1-14. Sample STDL0 Data Folder (TX16) 14
Figure 1-15. Containerization and Metadata 16
Figure 1-16. ESS Folder Organization 16
Figure 1-17. Sample STDL1 Data Folder (TX16) 17
Figure 1-18. STDL2 EEG Data Quality Report 18
Figure 1-19. Sample STDL2 Data Folder (TX16) 19
Figure 2-1. Vehicle Perturbation 29
Figure 2-2. Traffic Complexity Environment 33
Figure 2-3. Images of the Five Target Types for BCIT RVSP Experiments (X1 and X2) 37
Figure 2-4. Multiple Images of One Target Type for BCIT RVSP Experiments (X1 and X2) 37
Figure 2-5. Target Image Attributes for BCIT RVSP Experiments (X1 and X2) 38
Figure 2-6. BCIT RVSP (X1 and X2) EEG.event Data 39
Figure 2-7. BCIT RVSP Event Codes 39
Figure 2-8. BCIT RVSP Stimuli-Response Correlation 41
Figure 2-9. Screen Shot of Guard Duty Task (ID Card and Request for Access) 43
Figure 2-10. TX14 Data Collection System Architecture 54
Figure 2-11. TX14 Crewstation Interface 55
Figure 2-12. TX14 & TX15 Target Detection and Reporting 55
Figure 2-13. TX14 & TX15 Vehicle Intervention 56
Figure 2-14. TX15 Visual Oddball Task 60
Figure 2-15. TX16 Crewstation Interface 63
Figure 2-16. TX17A Simulated Roadway 66

Figure 2-17. RSVP CT2WS Target Images & Display Timing.. 73
Figure 2-18. RSVP CT2WS Central Cueing... 74
Figure 2-19. RSVP CT2WS Peripheral Cueing.. 75
Figure 2-20. RSVP Insurgent-Civilian Images 77
Figure 2-21. RSVP Insurgent-Civilian Targets.. 77
Figure 2-22. EEG Headsets and a Phantom Head 82
Figure 2-23. ODE System Architecture 93
Figure 2-24. ODE PRACTB Task 95
Figure 2-25. ODE Target Detection and Classification Task.. 96
Figure 2-26. RWN-VDE Experimental Procedures 101
Figure 2-27. A Bird's Eye View of the Event-Related Lane-Departure Paradigm 102
Figure 2-28. Reaction Time and Response Time Metrics 103
Figure 2-29. Event-Related Target Identification Paradigm 103
Figure 2-30. Steering Wheel 104
Figure 2-31. RWN-VDE Experimental Conditions 104
Figure 2-32. FLERP Monitoring Task 107
Figure 2-33. ACC Test Course.. 108
Figure 2-34. ACC Instrumented Vehicle.. 109
Figure 2-35. ACC Engineering Test Participant 110
Figure 2-36. Events Comprising an ACC Experimental Trial 110
Figure 2-37. Stock Interior and Exterior Views of the 2016 Ford Fusion Titanium 116
Figure 2-38. The Automotive Computer Shown Securely Mounted on the Rear Bench Seat (A), System Monitor Shown Securely Mounted to the Rear Headrest of the Passenger Seat (A, B) 117
Figure 2-39. Overview of System Developed to Support the Proposed Experiment................ 119
Figure 2-40. RWNVDEDP Timeline... 120
Figure 2-41, Exemplar stimuli for the Recognition Memory Task with the original (OLD) item on the left and the lure (NEW) item on the right in which only the age has been changed........... 127
Figure 2-42. Ride Motion Simulator.. 138
Figure 2-43. Control Screen for the Five PRACTB Tasks 140
Figure 2-44. Driving Course for ARL_ARPI_TX20 Experiments 142
Figure 2-45. ARL_ARPI_TX20 LFD Screen Presentation 143
Figure 2-46. Summary of the LFD paradigm for ARL_ARPI_TX20.. 144
Figure 2-47. Display Following a Collision with a Pedestrian for ARL_ARPI_TX20 146
Figure 2-48. Ride Motion Simulator.. 149
Figure 2-48. Driving Course for ARL_ARPI_TX22 Experiments 151
Figure 2-49. ARL_ARPI_TX22 LFD Screen Presentation 153
Figure 2-50. Summary of the LFD paradigm for ARL_ARPI_TX22.. 154

Figure 2-51. Display Following a Collision with a Pedestrian for ARL_ARPI_TX22 156
Figure 2-52. ARL_ARPI_TX22 Display Icons 156
Figure 2-54. Lane and Range Indicators... 158

List of Tables

Table 1. Revision History 1
Table 2. SANDR Acknowledgements 1
Table 3. ABDHS EEG Pipeline Tools 4
Table 4. SANDR Public Data Sets 21
Table 5. SANDR Public Data Metrics 21
Table 6. SANDR Instrumentation Summary 23
Table 7. SANDR Study Attribute Summary 24
Table 8. SANDR Repeat Subjects for BCIT Program Task 2 25
Table 9. SANDR Repeat Subjects for BCIT Program Task 3 25
Table 10. Driver and Passenger Subjects for RWNVDEDP Experiments 26

1 Data Repository Overview

The Standardized Annotated Neurophysiological Data Repository (SANDR) is a collection of standardized experimental data featuring test subject measurement via physiological instrumentation, with electroencephalographic (EEG) data as the central component across studies.

1.1 SANDR Summary Document

The purpose of this document is to provide information sufficient to download and work with repository data. The scope of the following dataset descriptions covers the SANDR datasets ready for use in large-scale data processing efforts.

In this document, descriptions of programs, studies and tasks were developed using the associated research protocols, design documents, and publications. There are multiple instances of wholesale "cut and paste" from source documents into this document, to provide insight into the characteristics of the research efforts and the resulting data.

1.1.1 Revision History

This document accounts for data ready to use in data analysis activities. As additional datasets are added to the repository the summary document will be updated accordingly, and version identifiers and release dates will be tracked via Table 1.

Table 1. Revision History

Title	Version	Date
ARL Experimental Data Set Summary	v1.0.0	NOV 2015
SANDR Data Summary	v2.0.0	SEP 2017
SANDR Data Summary	v2.0.1	NOV 2017
SANDR Data Description	v2.1.0	SEP 2018
SANDR Data Description	v2.1.1	MAR 2019
SANDR Data Description	v2.1.2	APR 2019
SANDR Data Description	v2.2.0	MAR 2021
SANDR Data Description	v2.2.1	SEP 2021
SANDR Data Description	V2.2.2	MAR 2023

1.1.2 Acknowledgements

SANDR development was sponsored by the U.S. Army Research Laboratory via the CaN CTA contract (number W911NF-10-2-0022). SANDR is populated with data resulting from joint ARL-GVSC (formerly TARDEC) studies, CaN CTA, and Autonomy Research Pilot Initiative (ARPI) studies. The following individuals are recognized for their contributions to SANDR through tool development, data curation, and data analysis:

Table 2. SANDR Acknowledgements

DEVCOM ARMY RESEARCH LABORATORY	Jonathan Touryan Anthony Ries	UTSA	Kay Robbins Jeremy Cockfield Kyung-min Su
dcs corp	Tony Johnson Bret Kellihan Courtney Crites Michael Dunkel Stephen Gordon Matthew Jaswa	intheon	Nima Bigdely-Shamlo Christian Kothe Tim Mullen
		UCSD	Scott Makeig Tzyy-Ping Jung

1.2 Data Repository Description

A goal of the CaN CTA Advanced Computational Approaches–SANDR (ACA-SANDR) project was to produce a large collection of well-annotated, synchronized data for shared analysis efforts. Moreover, given the complexity of real-world research, SANDR is intended to be a resource for supporting complex, large-scale analyses. Figure 1-1 illustrates the difference between the traditional path for data processing through publication, as compared with the resources added through the ACA-SANDR project, and in this example, analyzed as part of a big data analysis project, with EEG data called LARG.

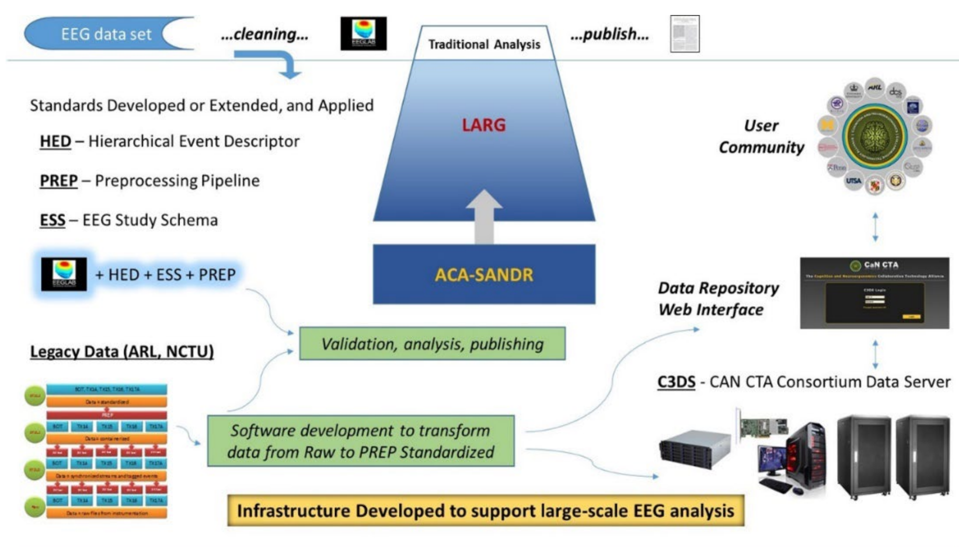


Figure 1-1. Traditional EEG Analysis Path versus Big Data Analysis

SANDR is currently hosted via the CaN CTA Consortium Data Server (C3DS), which features a user-friendly front-end for browsing and searching the repository. C3DS accounts must be approved by ARL, and will be established by the system administrator. The POC information is:

ARL Authorization for C3DS account
Jon Touryan
410-278-4329
jonathan.o.touryan.civ@mail.mil

C3DS System Administration
Bret Kellihan (DCS Corp.)
571-227-6284
bkellihan@dcscorp.com

1.2.1 Data Standardization Process

The SANDR datasets have been curated via a standardization process known as the "ARL Big Data for Human Sensing" (ABDHS) to support large-scale, cross-study analysis. The standards were adopted via collaborative efforts within the CaN CTA program, and applied to several "legacy" datasets that ARL produced in past research efforts that included neuroimaging. The pipeline steps are reflected in Figure 1-2.

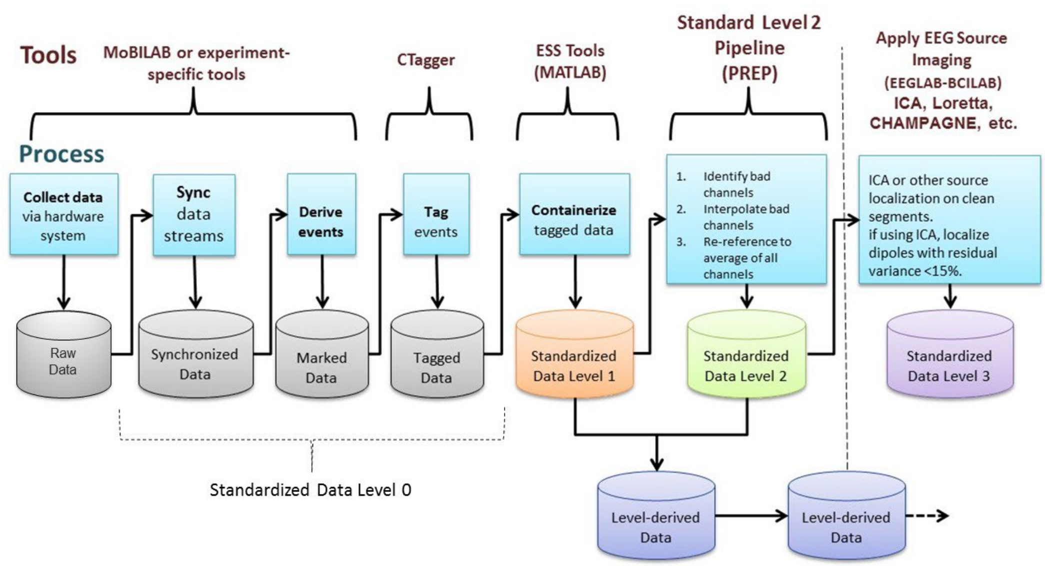


Figure 1-2. Big Data for Human Sensing EEG Pipeline Process

The ABDHS EEG pipeline is designed to manage the complexity involved with multi-lab and multi-study data analysis efforts by virtue of EEG data standardization. Through the application of common event nomenclature, a common EEG data structure, and common data organization schemes, automated pipeline tools can perform EEG data cleaning and validation checks to prepare the data for analysis. The pipeline is implemented through a set of standards and tools (identified in Table 3) that are applied to a study dataset in a series of stages (Raw, Level 0, Level 1, and Level 2), as shown in the system architecture depicted in Figure 1-3. The stages are described in the following subsections.

Table 3. ABDHS EEG Pipeline Tools

Tool Name	Definition	Availability
MATLAB	A multi-paradigm numerical computing environment	https://www.mathworks.com/products/matlab.html
EEGLAB	A MATLAB toolbox for processing EEG data and other physiological signals; defines the EEG data structure	https://sccn.ucsd.edu/eeglab/download.php
DART	The Data Annotation and Reprocessing Tool is a GUI-based MATLAB tool for processing Raw study data from a specific study	Formal request to the Army Research Laboratory
HED	Hierarchical Event Descriptor is a set of descriptor tags partially adopted from BrainMap/NeuroLex ontologies and organized hierarchically	http://www.hedtags.org/
ESS	EEG Study Schema is a standardized data specification and toolset for organizing EEG study data	http://www.eegstudy.org/
PREP	The PREP (preprocessing) Pipeline is a MATLAB-based tool that performs the following EEG data preprocessing: 1. Line noise removal 2. Identify bad channels 3. Interpolate bad channels 4. Re-reference to average of all channels	http://vislab.github.io/EEG-Clean-Tools/

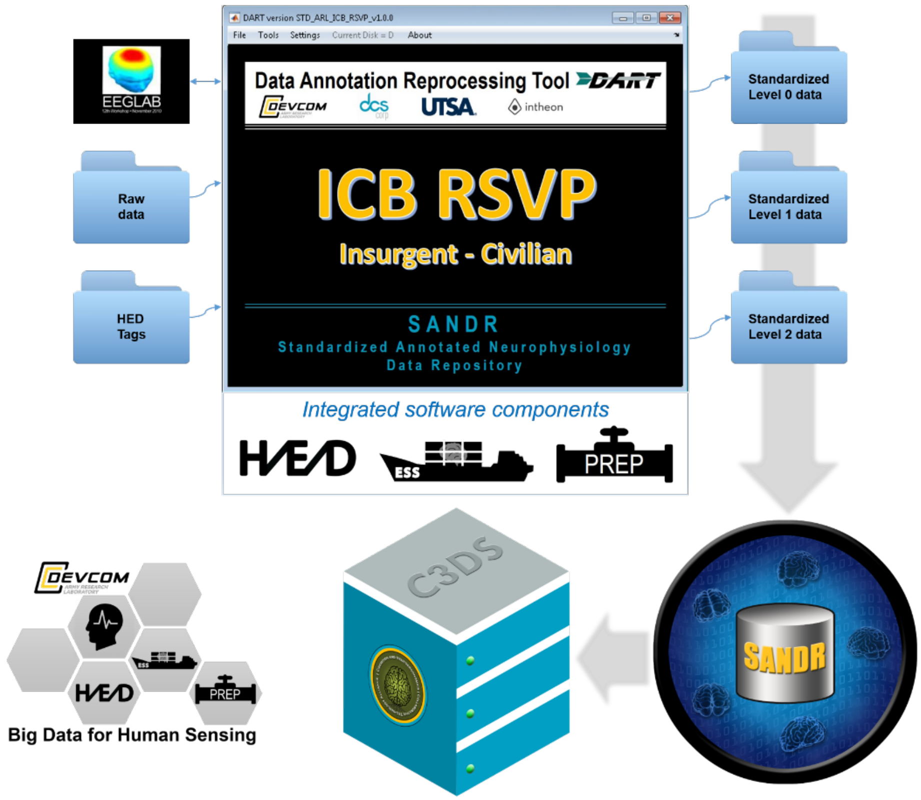


Figure 1-3. Big Data for Human Sensing System Architecture

1.2.1.1 Raw: Data directly from instrumentation

Raw Data is that which is transferred directly from the hardware platforms following data collection. It can also include files that required post-processing due to the nature of the collection method. For example, verbal communication that was captured via audio files and subsequently transcribed to produce representative text files may also be found within Raw data.

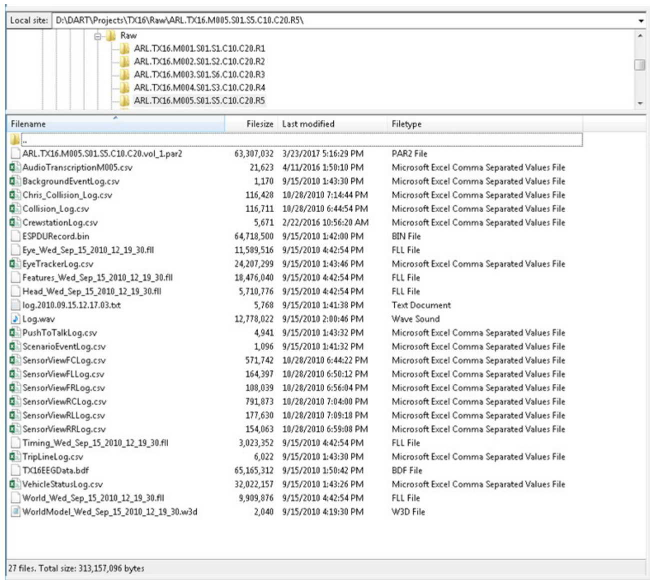


Figure 1-4. Sample Raw Data Folder (TX16)

Raw data can consist of numerous files for each recording session, or very few. The Raw dataset shown in Figure 1-4 is an example of a study that features numerous raw data files, which are a result of the complexity of the experimental objectives and the corresponding data collection system. Figure 1-5 depicts the system architecture used for collecting the TX16 experimental data.

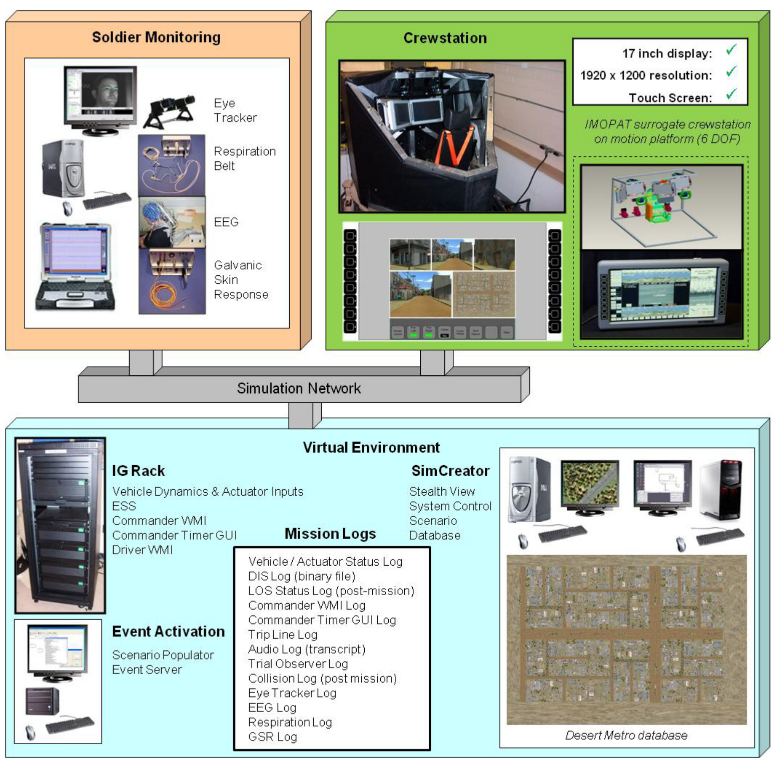


Figure 1-5. TX16 Data Collection System Architecture

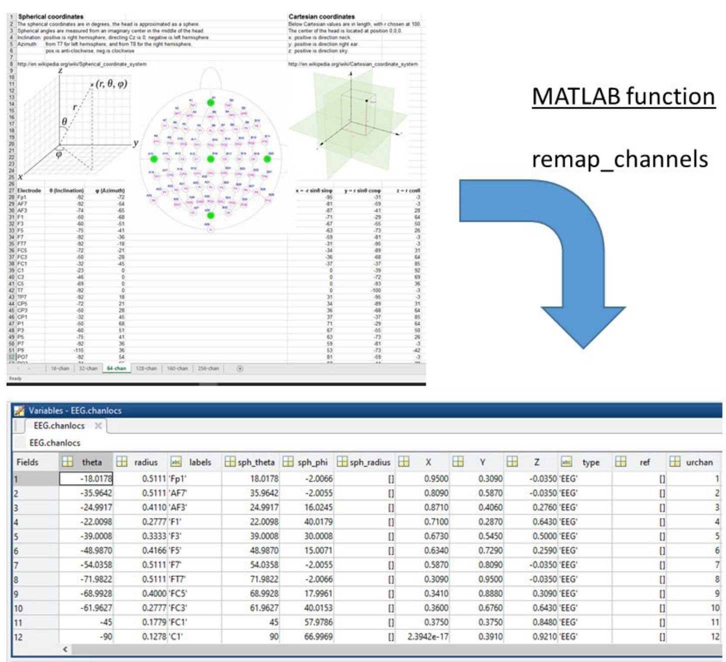
1.2.1.2 STDL0: Event Processing

The objective in converting Raw data to STDL0 data is to establish searchable discrete events of interest that can be used to interpret phenomena within the subject's physiological (e.g. EEG) data. Processing involves up to four main steps:

1) Insertion of EEG channel location data, if necessary
2) Synchronization of all data streams (achieved through common sync marker values logged on all data collection platforms), if necessary
3) Derivation of discrete events (which are marked with study-specific numeric event codes), if necessary
4) Application of a common vocabulary through event tag strings, which are based on the numeric event codes

1.2.1.2.1 EEG Channel Location Data

An essential process at STDL0 is to ensure that the EEG channel location data contains accurate coordinates based on the data collection equipment that was used. The BioSemi systems, for example, have specific channel coordinates for their caps based on the number of channels. The BioSemi cap specifications can be found at https://www.biosemi.com/headcap.htm.



Fields	theta	radius	labels	sph_theta	sph_phi	sph_radius	X	Y	Z	type	ref	urchan
1	-18.0178	0.5111	'Fp1'	18.0178	-2.0066	[]	0.9500	0.3090	-0.0350	'EEG'	[]	1
2	-35.9642	0.5111	'AF7'	35.9642	-2.0055	[]	0.8090	0.5870	-0.0350	'EEG'	[]	2
3	-24.9917	0.4110	'AF3'	24.9917	16.0245	[]	0.8710	0.4060	0.2760	'EEG'	[]	3
4	-22.0098	0.2777	'F1'	22.0098	40.0179	[]	0.7100	0.2870	0.6430	'EEG'	[]	4
5	-39.0008	0.3333	'F3'	39.0008	30.0008	[]	0.6730	0.5450	0.5000	'EEG'	[]	5
6	-48.9870	0.4166	'F5'	48.9870	15.0071	[]	0.6340	0.7290	0.2590	'EEG'	[]	6
7	-54.0358	0.5111	'F7'	54.0358	-2.0055	[]	0.5870	0.8090	-0.0350	'EEG'	[]	7
8	-71.9822	0.5111	'FT7'	71.9822	-2.0066	[]	0.3090	0.9500	-0.0350	'EEG'	[]	8
9	-68.9928	0.4000	'FC5'	68.9928	17.9961	[]	0.3410	0.8880	0.3090	'EEG'	[]	9
10	-61.9627	0.2777	'FC3'	61.9627	40.0153	[]	0.3600	0.6760	0.6430	'EEG'	[]	10
11	-45	0.1779	'FC1'	45	57.9786	[]	0.3750	0.3750	0.8480	'EEG'	[]	11
12	-90	0.1278	'C1'	90	66.9969	[]	2.3942e-17	0.3910	0.9210	'EEG'	[]	12

Figure 1-6. BioSemi Channel Location Remap

This information was developed by Dr. Kay Robbins at UTSA to develop a robust MATLAB algorithm for generating correct data for BioSemi EEG channel locations based on a function call from the STDL0 software, known as the Data Alignment and Reprocessing Tool (DART). After the call to "remap_channels", as shown in Figure 1-6, the EEG.chanlocs structure contains the information necessary to support STDL2 processing.

1.2.1.2.2 Data Synchronization

Some datasets are made available for inclusion in SANDR with data synchronization already completed, and numeric event codes written to the EEG.event structure. Other datasets require these steps to be performed as part of the STDL0 conversion.

The first synchronization step is to sync EEG data to the simulation data, and then sync eye-tracking data to the time-adjusted EEG data. Both sync processes use the same generic function that aligns data streams through a precise match of all sync markers that appear in both streams. From the subsequent time differentials, the following data elements are produced as function return values for use in synchronizing the two data stream: offset, drift, and jitter.

Time alignment is accomplished by applying the offset and the drift, bringing the first set of event times in line with the second. Figure 1-7 depicts a plot of sync markers from the EEG data stream in blue (y value = 2), and SIM data stream in red (y value = 1), with the x-axis representing time in seconds.

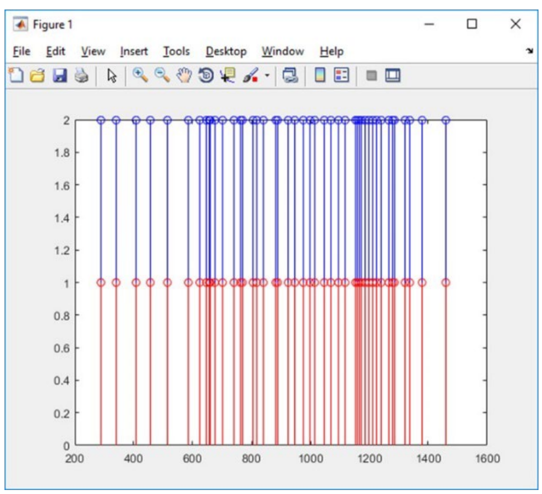


Figure 1-7. EEG to SIM Sync Plot (TX16)

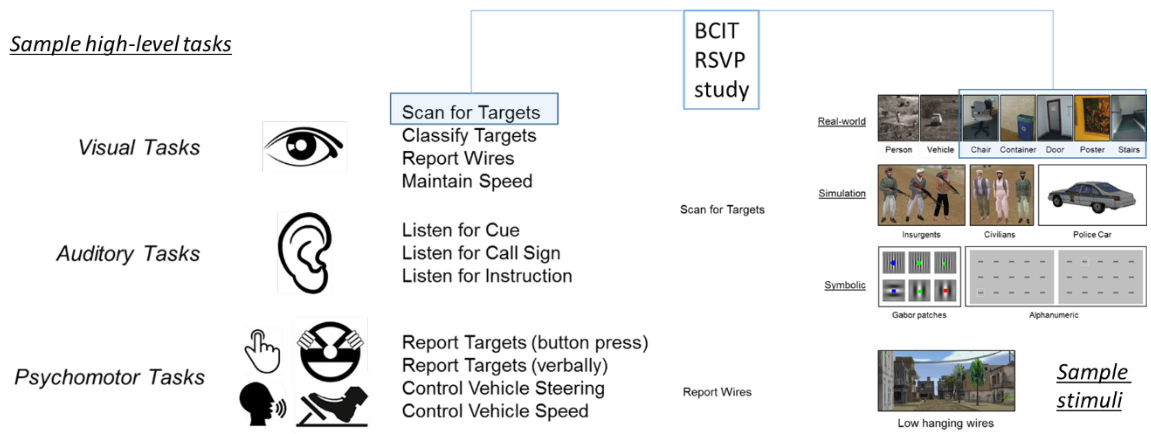
The derivation of events at STDL0 is sometimes necessary when performing calculations that are too computationally expensive to complete in real-time during data collection (for example, line-of-sight to targets), or when desiring to reference future events in order to more accurately tag present events (such as in perception tasks). For these datasets, the derived events are assigned numeric event codes, and written to a structure called EventData in the "_SIM" file. When reaching the point at which unique numeric event codes are associated with each event, and all events instances are marked with event codes, a dataset is ready to add text-based tags.

1.2.1.2.3 Event Derivation and Marking

Deriving events for a study requires examination of the research objectives, and specification of a series of event types that are based on the tasks performed by study participants. Figure 1-8 lists exemplar subject tasks from SANDR studies for the visual, auditory, and psychomotor modalities.

1.2.1.2.4 Task Tagging

When subjects perform experimental tasks, they are interacting with the research environment in some prescribed manner. For instance, in order to scan for targets, objects defined as targets must be presented to the subject for visual perception for a period beginning with onset and ending with offset. Note that the term "target" is study-specific. If the subject has also been instructed to report targets, they might do so by pressing a button, or by verbalizing the report.



Paradigm:
Paradigm/Oddball discrimination paradigm/Visual oddball paradigm/Rapid serial visual presentation

Attributes:
(Attribute/Condition/Target, Item/Object/Furniture/Chair, Sensory presentation/Visual, Participant/Effect/Visual, Participant/Effect/Cognitive/Target, Participant/Effect/Cognitive/Oddball),
(Attribute/Condition/Target, Item/Object/Furniture/Poster, Sensory presentation/Visual, Participant/Effect/Visual, Participant/Effect/Cognitive/Target, Participant/Effect/Cognitive/Oddball),
(Attribute/Condition/Target, Item/Object/Building/Door, Sensory presentation/Visual, Participant/Effect/Visual, Participant/Effect/Cognitive/Target, Participant/Effect/Cognitive/Oddball),
(Attribute/Condition/Target, Item/Object/Furniture/Container, Sensory presentation/Visual, Participant/Effect/Visual, Participant/Effect/Cognitive/Target, Participant/Effect/Cognitive/Oddball),
(Attribute/Condition/Target, Item/Object/Building/Stairs, Sensory presentation/Visual, Participant/Effect/Visual, Participant/Effect/Cognitive/Target, Participant/Effect/Cognitive/Oddball)

Figure 1-8. Task Paradigm and Tags

Task tags attempt to document items involved, actions taken, and senses used by subject in task, as well as the effect of the task (or event) on the subject.

1.2.1.2.5 Event Tagging

Text-based event descriptors are selected from the Hierarchical Event Descriptor (HED) library, or ontology (shown in Figure 1-9), which is a controlled vocabulary for tagging events. Multiple tags can be assigned to an event, each describing a specific aspect. The hierarchical format facilitates event searches at varying scopes, from the general to the very specific, in support of cross-study analysis efforts. Figure 1-10 illustrates events related to a vehicle perturbation in a driving study, along with the associated HED tags that specify several attributes about the event.

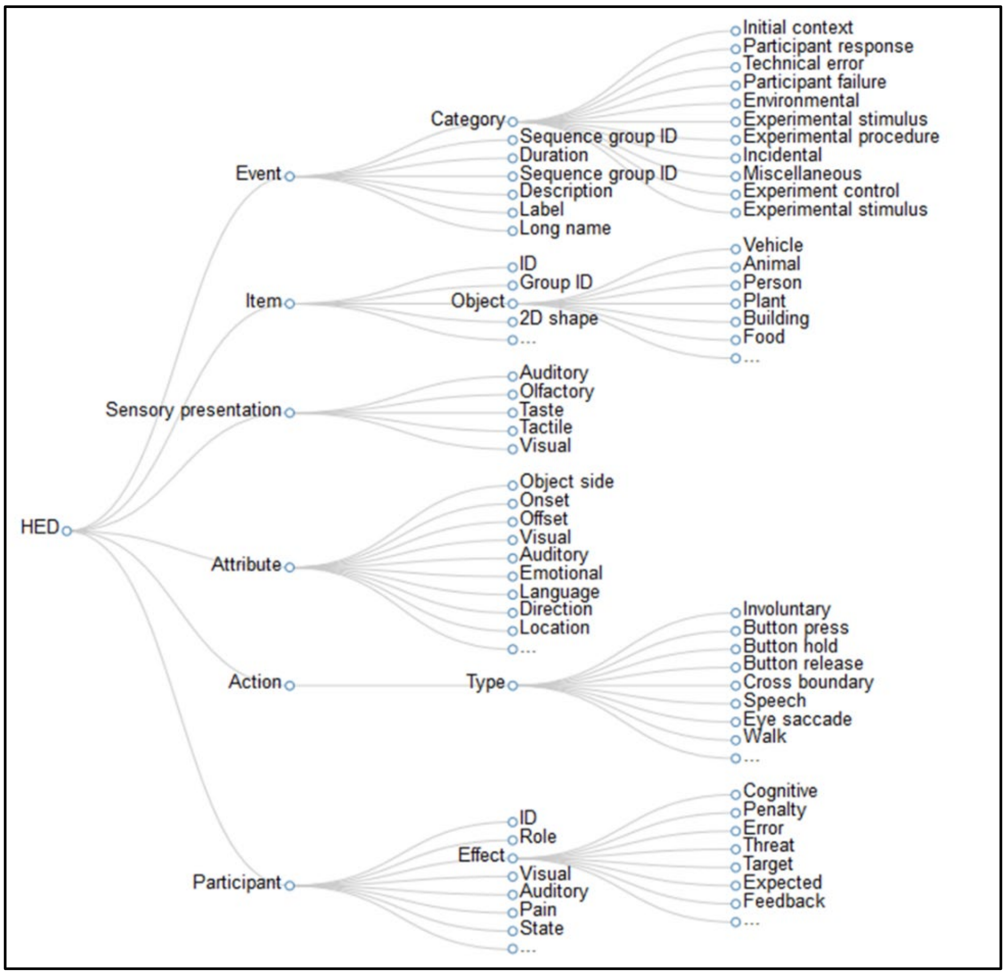
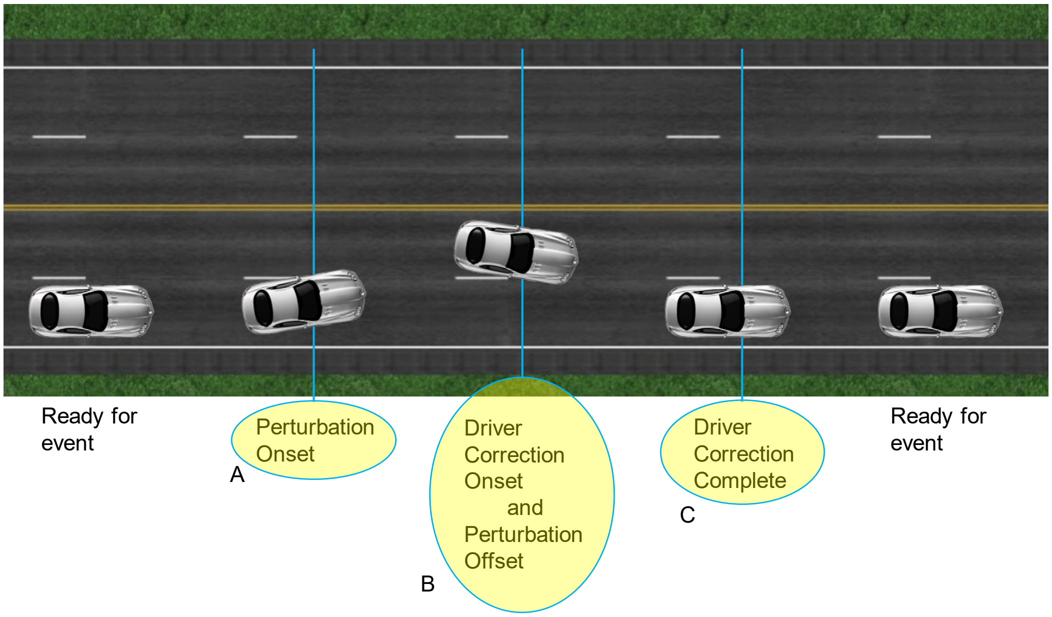


Figure 1-9. Hierarchical Event Descriptor

When the ontology is found to be lacking in the required breadth or depth to describe adequately the details of events, it can be expanded with new definitions through discussions with Dr. Kay Robbins (Kay.Robbins@utsa.edu) and Dr. Nima Bigdely-Shamlo (nima.bigdely@intheon.io).

Perturbation event

An increasing perpendicular force is applied from the left or right until the driver corrects



Associated HED Tags

A. <u>Perturbation Onset</u>
Event/Category/Experimental stimulus, Attribute/Onset, Event/Label/RightPerturbOnset, Event/Long name/Vehicle | Perturbation | Right | Onset, Event/Description/Beginning of a perturbation that moves the vehicle to the right via perpendicular force, (Item/Object/Vehicle/Car, Attribute/Object control/Perturb, Attribute/Direction/Right)

B. <u>Driver Correction Onset</u>
Event/Category/Participant response, Attribute/Onset, Event/Label/DriverCorrectOnset, Event/Long name/Behavioral | PerturbationResponse | Manual | Onset, Event/Description/Criteria for ending a perturbation achieved either ABS Heading Error more than 5.1566 degrees OR Driver is steering into the perturbation and STEER ANGLE more than 4 degrees, (Participant ~ Action/Control vehicle/Drive/Correct ~ Item/Object/Vehicle/Car)

B. <u>Perturbation Offset</u>
Event/Category/Experimental stimulus, Attribute/Offset, Event/Label/RightPerturbOffset, Event/Long name/Vehicle | Perturbation | Right | Offset, Event/Description/End of a perturbation that moves the vehicle to the right via perpendicular force, (Item/Object/Vehicle/Car, Attribute/Object control/Perturb, Attribute/Direction/Right)

C. <u>Driver Correction Offset</u>
Event/Category/Participant response, Attribute/Offset, Event/Label/DriverCorrectOffset, Event/Long name/Behavioral | PerturbationResponse | Manual | Offset, Event/Description/Criteria for ending a perturbation achieved either ABS Heading Error more than 5.1566 degrees OR Driver is steering into the perturbation and STEER ANGLE more than 4 degrees, (Participant ~ Action/Control vehicle/Drive/Correct ~ Item/Object/Vehicle/Car)

Figure 1-10. Sample HED Tags

Tags within HED can be browsed using an application called CTAGGER. A user interface (see Figure 1-11) provides access to the various levels of the hierarchy, and select attributes that correspond to events being tagged.

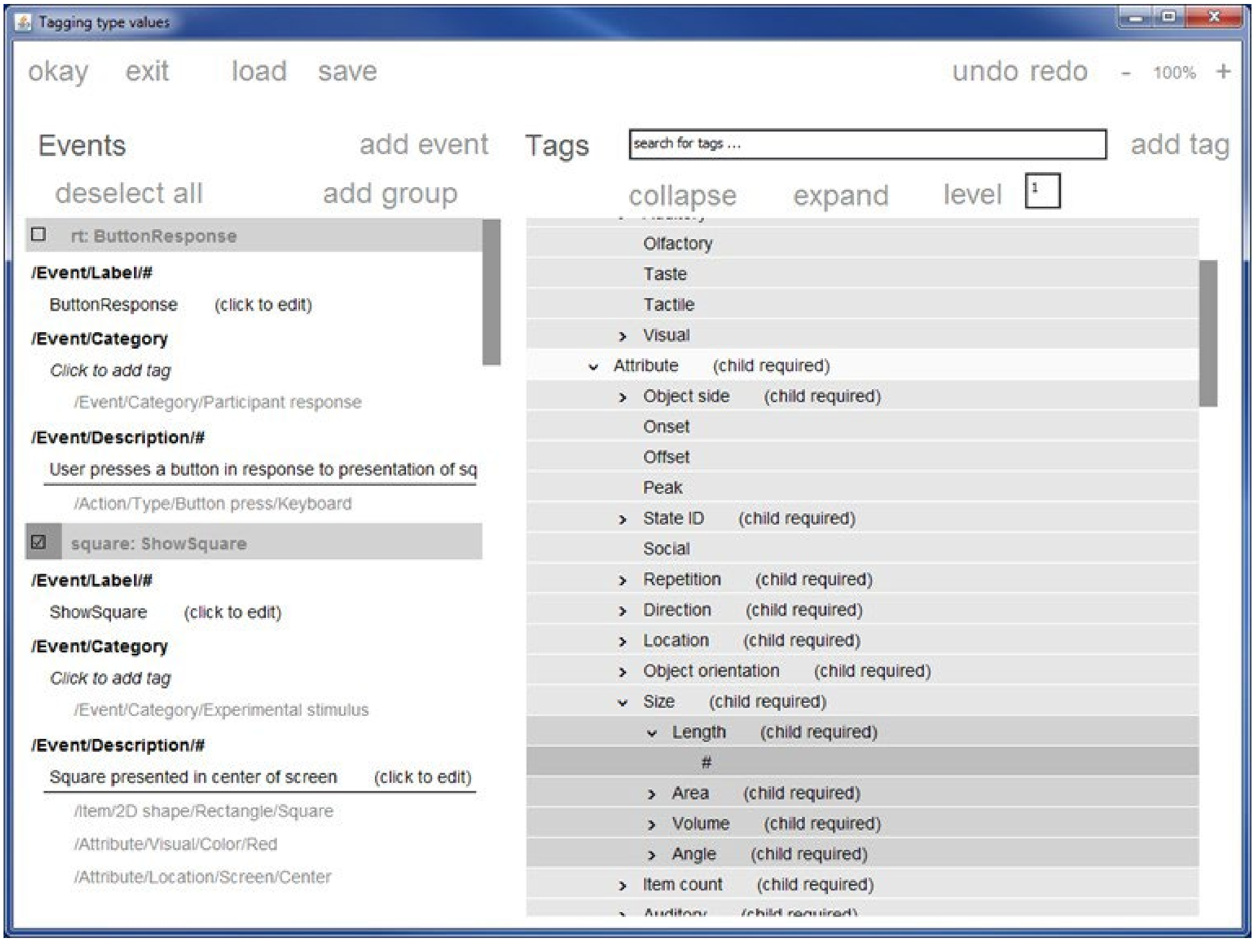


Figure 1-11. Sample HED Tag

Event tagging by the SANDR team at ARL involves identifying appropriate HED tags for each event, and then adding those tags to a spreadsheet that contains the numeric events codes as well. From that spreadsheet, a tag file is produced by a MATLAB script customized for the events that appear within each study. This file contains lookup data for each event, where the text of a tag is associated with the numeric event code.

Completed tag files can be validated using the online HED Validator, shown in Figure 1-12, to ensure that all of the event tags can be parsed and understood (i.e., are properly formatted), and utilize terminology defined in the selected HED version. The URL is http://netdb1.cs.utsa.edu/hed. As an alternative, validation may be accomplished using the MATLAB® function pop_tsvhed.m, included in the suite of HED tagging tools HEDTools-master, authored by Jeremy Cockfield at The University of Texas at Austin.

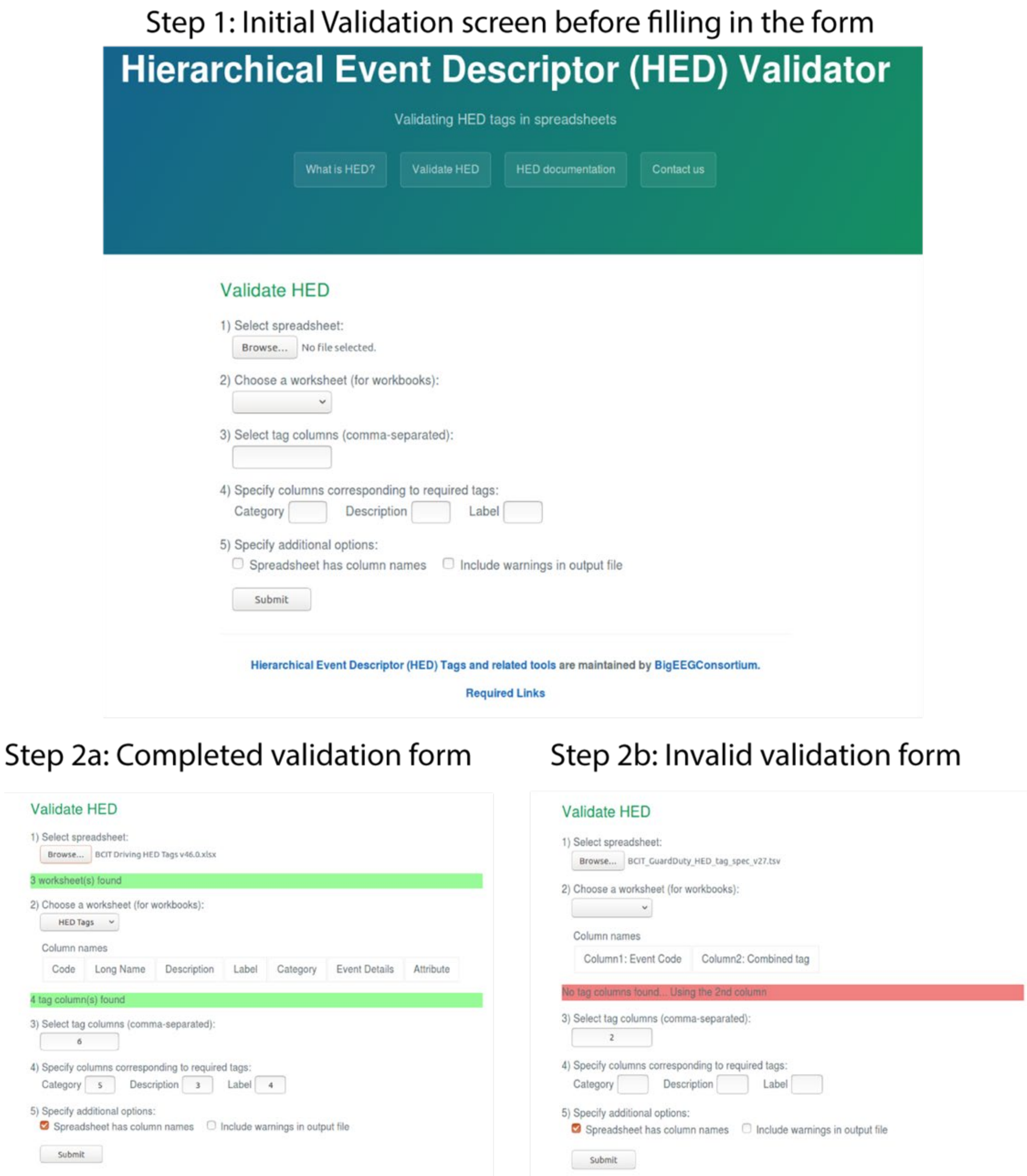


Figure 1-12. HED Validator

Once a HED tag file is validated, the HED tags can be applied by using DART, which contains processing to loop through events within the EEG.event structure, and uses the numeric event code as an index to look up the corresponding HED tag from the tag file. Each HED tag is then written to the corresponding event within the EEG.event structure, as shown in Figure 1-13.

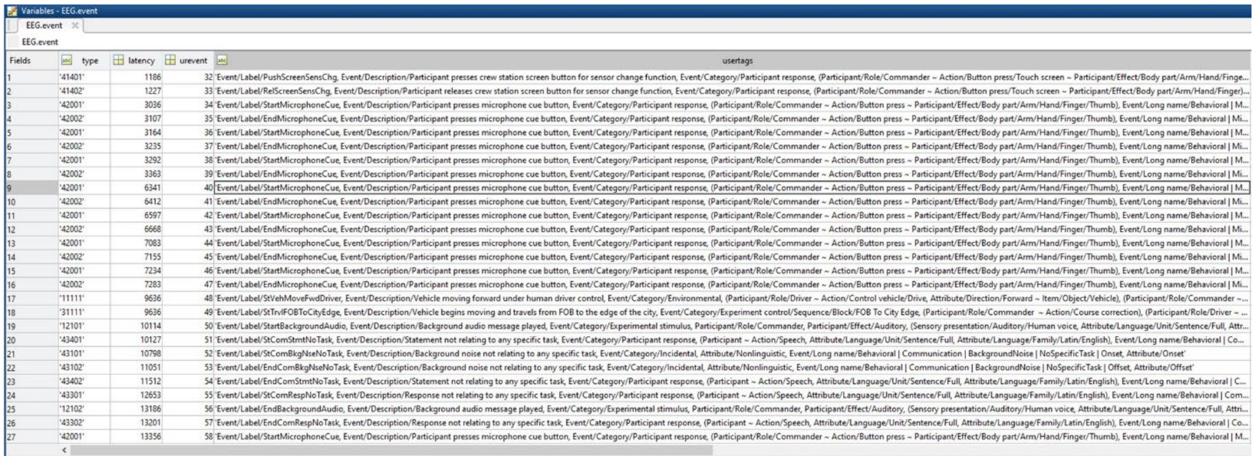


Figure 1-13. EEG.event Structure with HED Tags

At the end of the STDL0 process, the data folder contains a "*_EEG" file for every Raw dataset, and possibly a "*_SIM" and "*_EYE" file as well. Figure 1-14 shows the TX16 STDL0 folder, with "*_SIM.mat" and "*_EEG.set" files.

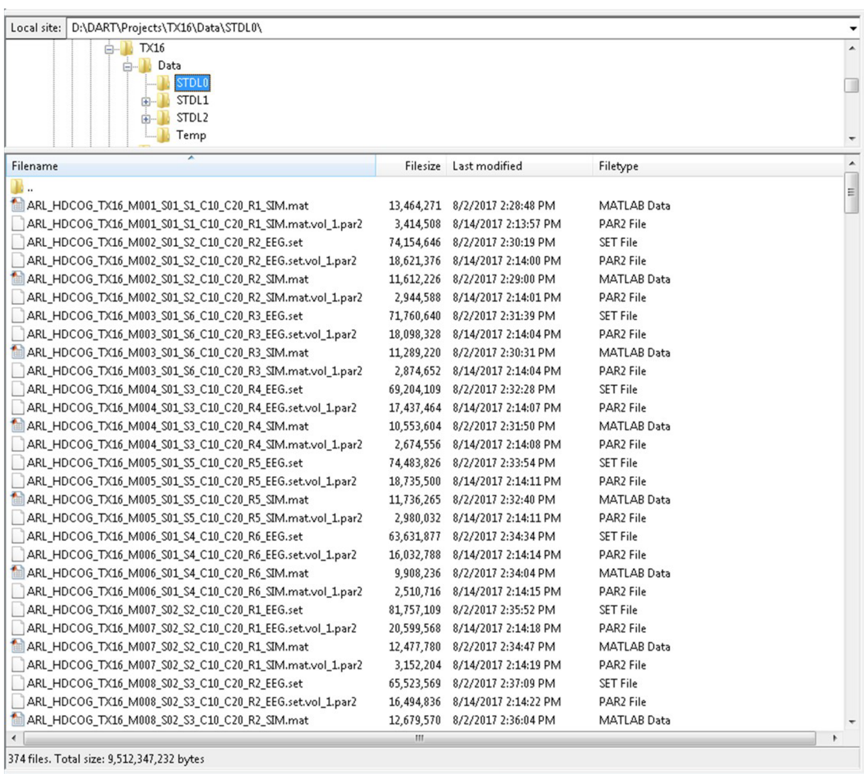


Figure 1-14. Sample STDL0 Data Folder (TX16)

Also shown are par2 files, which are produced for each STDL0 output file (EEG, SIM, and EYE). The par2 files provide checksum verification after a file transfer. Batch files that run via command window can be run to verify the clean transfer of MATLAB files, and will even regenerate the original file if the transferred copy is compromised. There are also MATLAB wrapper functions that have been developed to perform par2 generation and verification/regeneration by invoking the batch files using parameterized inputs. The use of par2 files can be especially helpful if it is not possible or convenient to download the data again from the C3DS.

1.2.1.3 STDL1: Containerization

The STDL0 files contain event-level tags, but more information about study-level aspects of the data needs to be added, and the data files need to be organized by subject (test participant) and session. The EEG Study Schema defines a container that characterizes EEG datasets in the form of an XML manifest (also generated as a JSON file).

STDL1 processing requires as input the set of STDL0 EEG files and study metadata that specifies the following: date and time information for all EEG files, study full description, study short description, study label(s), task label(s), task description(s), task paradigm(s), task attributes, experimenters, publications, funding organization, point of contact, project name (of the study), IRB protocol information, and copyright statement. Also processed at STDL1, if available, is the subject demographics data.

Data recording contents are organized according to a hierarchy, with the following levels (see Figure 1-15)

- Study
  A set of data recording sessions, collected to answer one or more related scientific questions.

- Session
  A single application of an EEG headset (cap on ⇨ cap off) for one or more subjects recorded within a single study. A session contains data from subjects performing one or more tasks.

- Task
  A task contains a single paradigm, and in combination, they allow answering scientific questions investigated in the study. Each paradigm features a set of related events.

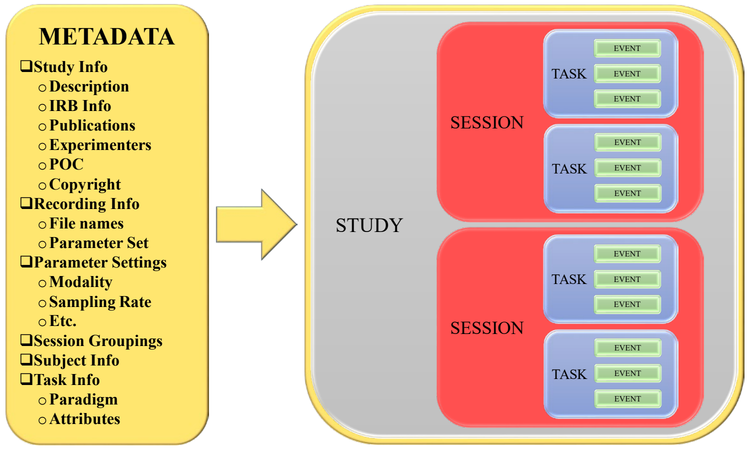


Figure 1-15. Containerization and Metadata

The container is generated as a level_1 folder with a series of session folders beneath, as shown in Figure 1-16. The session folders contain a set of EEG files for all tasks (runs) performed within the session by a subject, or a team of subjects being recorded simultaneously and working together on the same tasks. Each numbered folder contains the EEG recordings collected within a session, and the task name is embedded within the resulting EEG file name, which is modified when being copied from STDL0 to STDL1 (with no changes to the file contents).

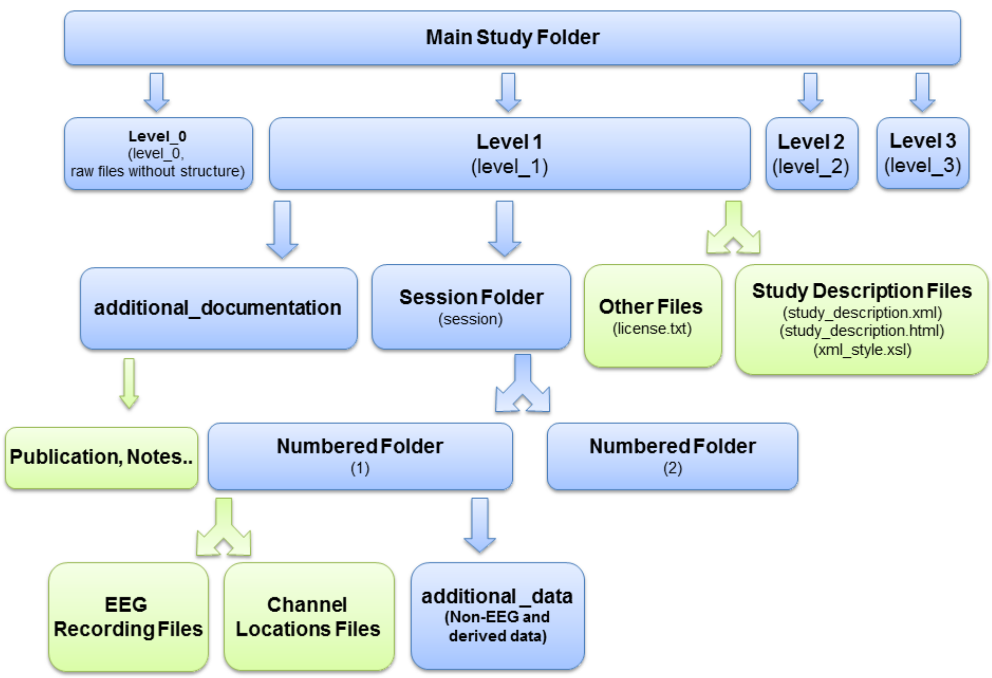


Figure 1-16. ESS Folder Organization

The study_description.xml file summarizes all of the data passed as STDL1 input, organized in an XML tree. Information such as recording_parameter_set, for example, is derived during the STDL1 conversion process, providing groups of related information such as number of EEG channels, sampling rate, and modalities included in the channel data. Each data recording, organized by session, is then linked with a specific recording_parameter_set, establishing a relationship between data and metadata. Figure 1-17 represents the view of a session, and associated folder contents, within a STDL1 folder.

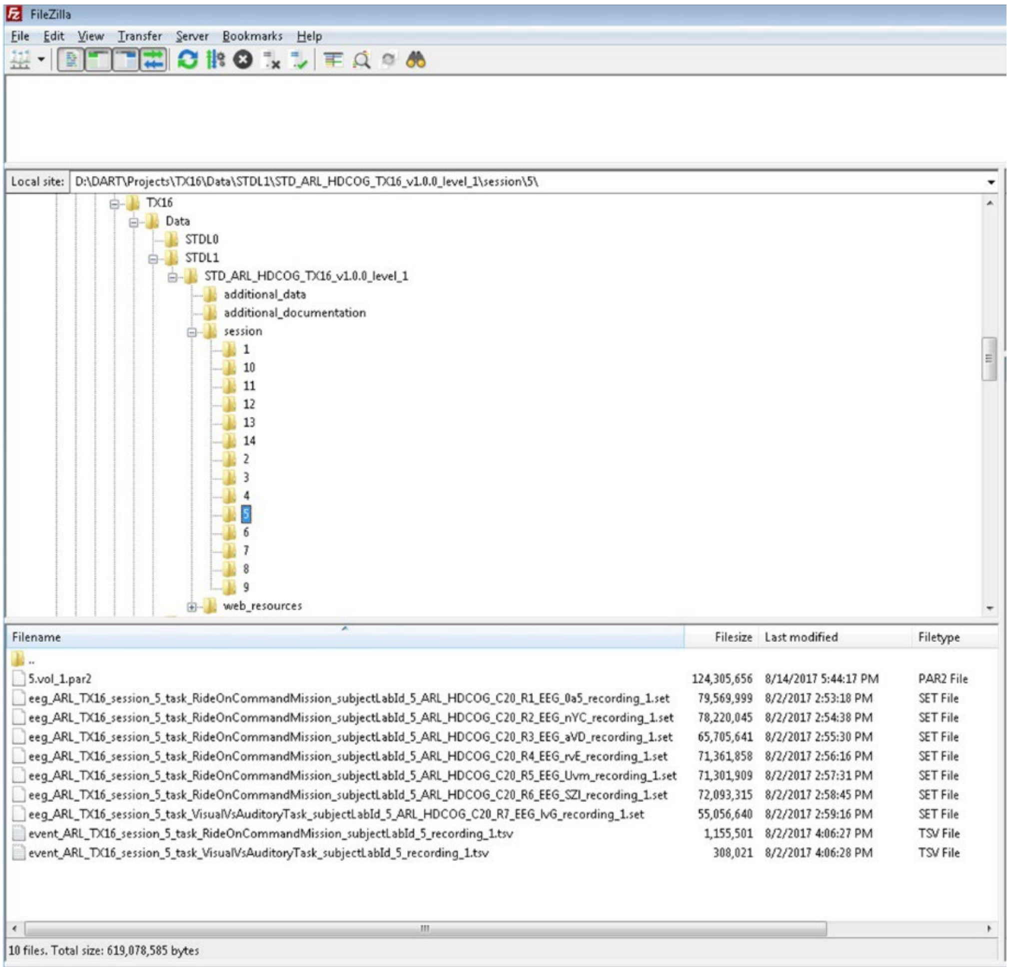


Figure 1-17. Sample STDL1 Data Folder (TX16)

1.2.1.4 STDL2: Re-referencing and bad channel detection

STDL2 processing provides standardized robust referencing, line noise removal, and bad channel detection. Furthermore, bad channel data are removed, and the contents replaced with interpolated data from neighboring electrodes. Note that the size of each STDL2 EEG file is significantly greater than its STDL1 counterpart. This is because EEG channel data are saved at STDL2 using double precision floats instead of single precision, due to the extensive calculations performed at this level.

Comprehensive data quality reports are generated at STDL2 as pdf files, and placed in the session folder along with each EEG file. The contents of the report are summarized in Figure 1-18.

Visualize the EEG output from the PREP processing pipeline.

Table of Contents

Write data status and report header 2
Line noise removal step 2
Initial detrend for reference calculation 2
Spectrum after line noise and detrend 3
Referencing step 5
Robust channel deviation (referenced) 6
Robust channel deviation (original) 8
Robust channel deviation (interpolated) 9
Robust deviation window statistics 9
Median max abs correlation (referenced) 11
Median max abs correlation (original) 12
Median max abs correlation (interpolated) 13
Mean max abs correlation (referenced) 14
Mean max abs correlation (original) 15
Mean max abs correlation (interpolated) 16
Bad min max correlation fraction (referenced) 17
Bad min max correlation fraction(original) 18
Bad min max correlation fraction (interpolated) 19
Correlation window statistics 19
Bad ransac fraction (referenced) 21
Bad ransac fraction (original) 22
Bad ransac fraction (interpolated) 23
Channels with poor ransac correlations 23
HF noise Z-score (referenced) 25
HF noise Z-score (original) 26
HF noise Z-score (interpolated) 27
HF noise window stats 27
Noisy average vs robust average reference 29
Noisy and robust average reference by time 30
Noisy vs robust average reference (filtered) 30
Noisy minus robust average reference by time 32

Calling directly: prepReport

This helper reporting script expects that EEGReporting will be in the base workspace with an EEGReporting.etc.noiseDetection structure containing the report. It also expects the following variables in the base workspace:

- summaryFile - variable containing the open file descriptor for summary
- consoleID - variable with open file descriptor for console (usually 1 unless the output is redirected).
- relativeReportLocation report location relative to summary

The reporting function appends a summary to the summary report.

1

Figure 1-18. STDL2 EEG Data Quality Report

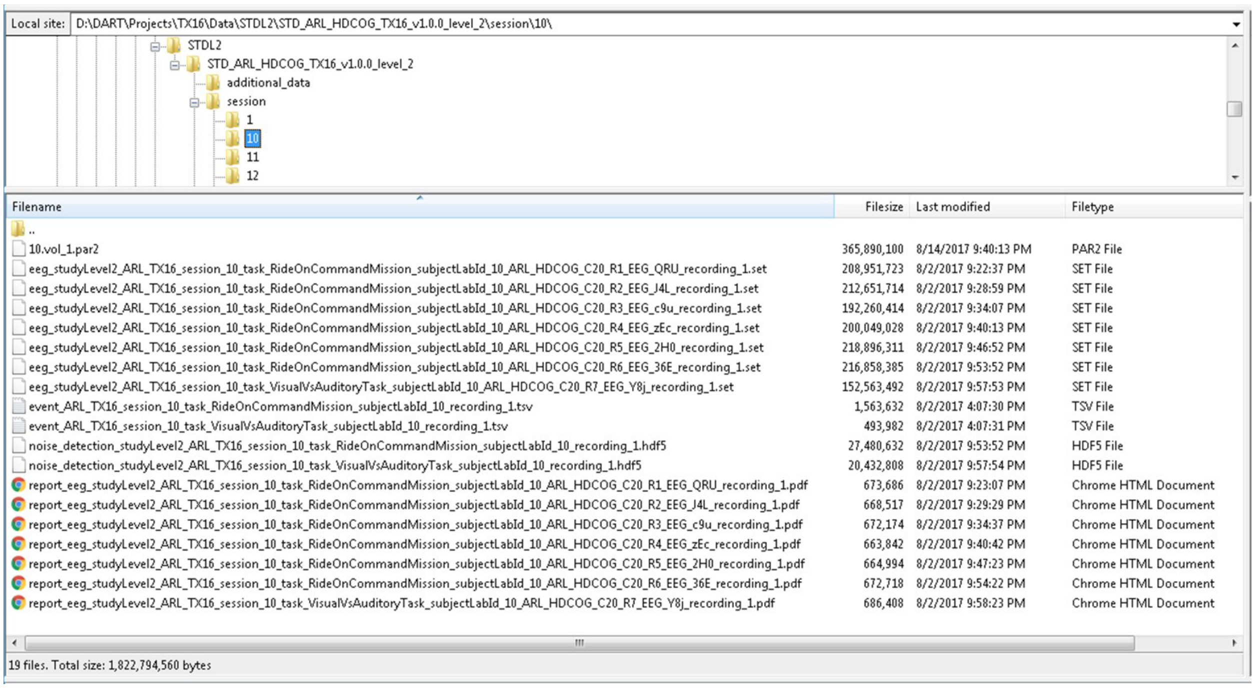


Figure 1-19. Sample STDL2 Data Folder (TX16)

The STDL2 data folders are very similar to the structure of the STDL1 folders, although there is additional information at STDL2, as can be seen in Figure 1-19.

1.2.1.5 STDL2 256: Down-sampled STDL2 EEG (256Hz)

A processing stage that occurs outside the formal pipeline is one that supports the ease of use of SANDR data. It is advantageous to have a replica set of all SANDR EEG files with a 256 Hz sampling rate to facilitate faster downloads and processing times when testing analysis efforts or building machine learning models. Thus, it is a goal of the SANDR project to generate an additional down-sampled 256 Hz EEG file for all files at STDL2 if the native file is recorded at a higher sampling rate.

1.2.1.6 Data transfer and par2

When copying EEG files from one disk to another it is possible for files to become corrupted in the process. To help prevent data loss, each dataset has associated par2 files that can be used either to verify files (or folders) after copy, or to rebuild files (or folders) if they are corrupted.

The tools that perform the verification and/or file or folder rebuild are on the C3DS, at **/Tools/par2/Batch programs** and **/Tools/par2/MATLAB functions**. Usage is provided via internal documentation (i.e., comments in the source files).

1.3 SANDR-related Publications

Thomas Rognon, Rebecca Strautman, Lauren Jett, Nima Bigdely-Shamlo, Scott Makieg, Tony Johnson, Kay A. Robbins (2013) **CTAGGER: Semi-structured community tagging for annotation and data-mining in event-rich contexts**, *Conference: 2013 IEEE Global Conference on Signal and Information Processing (GlobalSIP)* December 2013
DOI: 10.1109/GlobalSIP.2013.6736797
http://ieeexplore.ieee.org/document/6736797/

Nima Bigdely-Shamlo, Tim Mullen, Christian Kothe, Kyung-Min Su, and Kay A. Robbins (2015) **The PREP pipeline: standardized preprocessing for large-scale EEG analysis**, *Frontiers in Neuroinformatics* 9 (2015)
DOI: doi.org/10.3389/fninf.2015.00016
https://doi.org/10.3389/fninf.2015.00016

Nima Bigdely-Shamlo, Jeremy Cockfield, Scott Kerick, Thomas Rognon, Chris LaValle, Makato Miyakoshi, and Kay A. Robbins (2016) **Hierarchical Event Descriptors (HED): Semi-Structured Tagging for Real-World Events in Large-Scale EEG**, *Frontiers in Neuroinformatics* 10:42
DOI: 10.3389/fninf.2016.00042
https://doi.org/10.3389/fninf.2016.00042

Nima Bigdely-Shamlo, Scott Makeig, and Kay A. Robbins (2016) **Preparing Laboratory and Real-World EEG Data for Large-Scale Analysis: A Containerized Approach**, *Frontiers in Neuroinformatics* 10 (2016)
DOI: doi.org/10.3389/fninf.2016.00007
https://doi.org/10.3389/fninf.2016.00007

Kelly Kleifges, Nima Bigdely-Shamlo, Scott Kerick, and Kay Robbins (2017) "**BLINKER: Automated extraction of ocular indices from EEG enabling large-scale analysis**, *Frontiers in Neuroscience* 11 (FEB 2017)
DOI: 10.3389/fnins.2017.00012
https://doi.org/10.3389/fnins.2017.00012

1.4 Summary of Datasets

A list of the publicly available data currently standardized and archived in the repository, including dataset metrics, are provided in Table 4 and Table 5.

Table 4. SANDR Public Data Sets

SANDR Data Sets											
Experimental Data Set by Organization		Experimental Data Totals			Standardized Level					Public	C3DS Server
		Subjects	:	Data sets	Raw	0	1	2	2 (256Hz)	Access	Location
ARL_BCIT_CalibrationDriving	BCIT T1/T2/T3 XC	206	:	247	1	1	1	1	1	1	/ARL_BCIT
ARL_BCIT_BaselineDriving	BCIT T1/T2/T3 XB	128	:	128	1	1	1	1	1	1	/ARL_BCIT
ARL_BCIT_TrafficComplexity	BCIT T2 X2	29	:	30	1	1	1	1	1	1	/ARL_BCIT
ARL_BCIT_SpeedControl	BCIT T2 X6	32	:	63	1	1	1	1	1	1	/ARL_BCIT
ARL_BCIT_AuditoryCueing	BCIT T2 X7	17	:	34	1	1	1	1	1	1	/ARL_BCIT
ARL_BCIT_MindWandering	BCIT T2 X8	21	:	60	1	1	1	1	1	1	/ARL_BCIT
ARL_BCIT_RSVPBaseline	BCIT T3 X1	27	:	27	1	1	1	1	1	1	/ARL_BCIT
ARL_BCIT_RSVPExpertise	BCIT T3 X2	10	:	51	1	1	1	1	1	1	/ARL_BCIT
ARL_BCIT_BasicGuardDuty	BCIT T3 X3	21	:	21	1	1	1	1	1	1	/ARL_BCIT
ARL_BCIT_AdvancedGuardDuty	BCIT T3 X4	27	:	27	1	1	1	1	1	1	/ARL_BCIT
ARL_CANCTA_RWNVDEDP	Real-world driving study (dyads)	44	:	43	1	1	1	1	1	1	/ARL_CANCTA_RWNVDEDP
ARL_HDCOG_TX14	Target Detection/UGV Control	20	:	147	1	1	1	1	1	1	/ARL_HDCOG_TX14
ARL_HDCOG_TX15	Target Detection/UGV Control	14	:	94	1	1	1	1	1	1	/ARL_HDCOG_TX15
ARL_HDCOG_TX15Oddball	Oddball (Gabor patch)	8	:	8	1	1	1	1	1	1	/ARL_HDCOG_TX15
ARL_HDCOG_TX16	Vehicle Commander multi-tasking	14	:	81	1	1	1	1	1	1	/ARL_HDCOG_TX16
ARL_HDCOG_TX16AuditoryVisual	Auditory vs Visual task (Scenario 7)	13	:	13	1	1	1	1	1	1	/ARL_HDCOG_TX16
ARL_HDCOG_TX17A	Driver Fatigue / racetrack	13	:	20	1	1	1	1	1	1	/ARL_HDCOG_TX17A
ARL_ARPI_TX20	Measures of Trust in Automation	24	:	119	1	1	1	1	1	1	/ARL_ARPI_TX20
ARL_ARPI_TX22	Measures of Trust in Automation	17	:	68	1	1	1	1	1	1	/ARL_ARPI_TX22
ARL_ICB_CT2WS	RSVP CT2WS	17	:	72	1	1	1	1	1	1	/ARL_ICB_CT2WS
ARL_ICB_RSVP	RSVP Insurgent-Civilian	16	:	51	1	1	1	1	1	1	/ARL_ICB_RSVP
ARL_EEGCS_VEP	From EEG Comparison study	18	:	18		1	1	1	1	1	/ARL_VEP
DCS_CANCTA_FT	Finger Tapping	14	:	14	1	1	1	1	1	1	/DCS_CANCTA_FT
DCS_CANCTA_ODE	Operator Dynamics of Event Appraisal	17	:	67	1	1	1	1	1	1	/DCS_CANCTA_ODE
NCTU_CANCTA_RWN_VDE	DSS+PVT+MD+DF	17	:	855		1	1	1	1	1	/NCTU_CANCTA_RWN_VDE
TNO_CANCTA_ACC	Adaptive Cruise Control	21	:	45	1	1	1	1	1	1	/TNO_CANCTA_ACC
TNO_CANCTA_FLERP	Fixation-locked ERP	15	:	42	1	1	1	1	1	1	/TNO_CANCTA_FLERP
27 studies		**820**		**2,445**	**25**	**27**	**27**	**27**	**27**	**27**	https://dev.cancta.net/C3DS

Table 5. SANDR Public Data Metrics

SANDR Public Data Metrics										
Experimental Data Set by Organization	STDL2 Data Num Subjects	:	STDL2 Data Num Sessions	Number of Datasets	Hours of EEG	Data Size (GB)	Number of Event Types	Number of Zero-instance Events Types	Number of Event Types Used	Number of Total Event Instances
ARL_BCIT_CalibrationDriving	206	:	247	247	63.56	380.80	123	103	20	97,971
ARL_BCIT_BaselineDriving	128	:	128	128	131.47	883.40	123	95	28	168,288
ARL_BCIT_TrafficComplexity	29	:	29	29	22.87	50.63	123	86	37	30,737
ARL_BCIT_SpeedControl	32	:	32	63	44.76	99.32	246	227	19	108,594
ARL_BCIT_AuditoryCueing	17	:	17	34	26.00	57.85	246	221	25	102,524
ARL_BCIT_MindWandering	21	:	21	60	30.05	69.26	369	326	43	80,792
ARL_BCIT_RSVPBaseline	27	:	27	27	31.83	271.20	46	11	35	517,597
ARL_BCIT_RSVPExpertise	10	:	51	51	59.72	493.10	230	196	34	1,014,929
ARL_BCIT_BasicGuardDuty	21	:	21	21	19.07	165.00	40	17	23	20,270
ARL_BCIT_AdvancedGuardDuty	27	:	27	27	23.87	191.10	40	11	29	27,285
ARL_CANCTA_RWNVDEDP	44	:	43	43	129.77	33.85	406	98	308	346,679
ARL_HDCOG_TX14	20	:	20	147	28.42	15.52	368	254	114	55,254
ARL_HDCOG_TX15	14	:	14	94	18.68	10.38	736	617	119	44,358
ARL_HDCOG_TX15Oddball	8	:	8	8	1.48	0.82				
ARL_HDCOG_TX16	14	:	14	81	26.16	14.52	486	269	217	280,015
ARL_HDCOG_TX16AuditoryVisual	13	:	13	13	3.49	1.94				
ARL_HDCOG_TX17A	13	:	13	20	15.95	10.94	18	5	13	12,271
ARL_ARPI_TX20	24	:	24	119	26.32	85.26	516	387	129	78,045
ARL_ARPI_TX22	17	:	17	68	19.70	46.95	258	150	108	102,348
ARL_ICB_CT2WS	17	:	18	72	17.23	7.97	192	72	120	51,190
ARL_ICB_RSVP	16	:	16	51	6.39	39.77	192	91	101	136,234
ARL_EEGCS_VEP	18	:	18	18	2.82	3.26	7		7	10,322
DCS_CANCTA_FT	14	:	14	14	17.83	27.65	91	6	85	29,190
DCS_CANCTA_ODE	17	:	17	67	18.74	15.30	186	116	70	10,686
NCTU_CANCTA_RWN_VDE	17	:	855	855	207.62	129.80	79		79	364,735
TNO_CANCTA_ACC	21	:	15	45	20.50	83.40	177	70	107	70,120
TNO_CANCTA_FLERP	15	:	21	42	23.02	16.56	45	5	40	159,508
27	**820**		**1,740**	**2,444**	**1,037.34**	**3,205.55**	**5,343**	**3,433**	**1,910**	**3,919,942**

SANDR data includes EEG data, and other modalities as needed for each research objective, as listed in Table 6.

Table 6. SANDR Instrumentation Summary

SANDR Public Data Instrumentation										
Experimental Data Set by Organization	EEG System Type	EEG Num Channels	EEG Sampling Rate	EYE-tracking System Type	EYE-tracking Sampling Rate	EMG System Type	ECG System Type	HR System Type	EDA System Type	Actigraph System Type
ARL_BCIT_CalibrationDriving	BioSemi	64 & 256	1024 & 2048	SMI RED250	250 Hz					
ARL_BCIT_BaselineDriving	BioSemi	64 & 256	1024 & 2048	SMI RED250	250 Hz					
ARL_BCIT_TrafficComplexity	BioSemi	64	1024	SMI RED250	250 Hz					
ARL_BCIT_SpeedControl	BioSemi	64	1024	SMI RED250	250 Hz					
ARL_BCIT_AuditoryCueing	BioSemi	64	1024	SMI RED250	250 Hz					
ARL_BCIT_MindWandering	BioSemi	64	1024	SMI RED250	250 Hz					
ARL_BCIT_RSVPBaseline	BioSemi	256	1024	SMI RED250	250 Hz					
ARL_BCIT_RSVPExpertise	BioSemi	256	1024	SMI RED250	250 Hz					
ARL_BCIT_BasicGuardDuty	BioSemi	256	1024	SMI RED250	250 Hz					
ARL_BCIT_AdvancedGuardDuty	BioSemi	256	1024	SMI RED250	250 Hz					
ARL_CANCTA_RWNVDEDP	ABM	24	256				Zephyr Bioharness 3	Zephyr Bioharness 3	Empatica E4	Readiband
ARL_HDCOG_TX14	BioSemi	64	256	faceLAB4	60 Hz					
ARL_HDCOG_TX15	BioSemi	64	256	faceLAB4	60 Hz					
ARL_HDCOG_TX15Oddball	BioSemi	64	256	faceLAB4	60 Hz					
ARL_HDCOG_TX16	BioSemi	64	256	faceLAB4	60 Hz					
ARL_HDCOG_TX16AuditoryVisual	BioSemi	64	256	faceLAB4	60 Hz					
ARL_HDCOG_TX17A	BioSemi	64	256	SmartEye	60 Hz					
ARL_ARPI_TX20	BioSemi	64	1024 & 8192	SmartEye	60 Hz		BioSemi		BioSemi	
ARL_ARPI_TX22	BioSemi	64	1024	SmartEye	60 Hz		BioSemi		BioSemi	
ARL_ICB_CT2WS	BioSemi	64	512							
ARL_ICB_RSVP	BioSemi	64	1024							
ARL_EEGCS_VEP	BioSemi	64	1024							
DCS_CANCTA_FT	BioSemi	256	1024							
DCS_CANCTA_ODE	BioSemi	64	1024	faceLAB4		BioSemi	BioSemi		BioSemi	
NCTU_CANCTA_RWN_VDE	Neuro Scan	64	1000	SMI RED			Neuro Scan			Readiband
TNO_CANCTA_ACC	BioSemi	64	2048			BioSemi	BioSemi	BioSemi		
TNO_CANCTA_FLERP	BioSemi	32	512	SmartEye						

SANDR data also includes a variety of attributes that are associated with the experimental tasks and procedures. Some of the datasets overlap with each other due to common paradigms, such as RSVP experiments, or common events, such as the various driving experiments. Table 7 represents a matrix where attributes that apply to a given data set are shown with a green "1".

Table 7. SANDR Study Attribute Summary

SANDR Study Attributes	Real WoRLD	SIM	Under Motion	Driving	Autonomous Steering Control	Autonomous Speed Control	Human Driver not under Study	Navigation Responsibility	Perturbations	Interactions with Humans	Audio Played	Interactions with Technology	Target Detection	Human Targets	Human Non-targets	Vehicle Targets	Vehicle Non-targets	Naturalistic Object Targets	Naturalistic Object Non-targets	Geographic Shape/pattern Targets	Geographic Shape/pattern Non-targets	RSVP Paradigm	Oddball Paradigm
ARL_BCIT_CalibrationDriving	0	1	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ARL_BCIT_BaselineDriving	0	1	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
ARL_BCIT_TrafficComplexity	0	1	0	1	0	0	0	1	1	0	0	0	0	0	1	0	1	0	1	0	0	0	0
ARL_BCIT_SpeedControl	0	1	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
ARL_BCIT_AuditoryCueing	0	1	0	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0
ARL_BCIT_MindWandering	0	1	0	1	0	0	0	1	1	0	1	0	1	0	0	1	1	0	1	0	0	0	0
ARL_BCIT_RSVPBaseline	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0
ARL_BCIT_RSVPExpertise	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	1	0
ARL_BCIT_BasicGuardDuty	0	1	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
ARL_BCIT_AdvancedGuardDuty	0	1	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0
ARL_CANCTA_RWNVDEDP	1	0	1	1	0	0	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0
ARL_HDCOG_TX14	0	1	1	1	1	1	0	0	0	0	0	1	1	1	0	0	1	0	1	0	0	0	0
ARL_HDCOG_TX15	0	1	1	1	1	1	0	0	0	0	0	1	1	1	0	0	1	0	1	0	0	0	0
ARL_HDCOG_TX15Oddball	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	1
ARL_HDCOG_TX16	0	1	1	0	0	0	1	1	0	1	1	1	1	1	1	0	1	1	1	0	0	0	0
ARL_HDCOG_TX16AuditoryVisual	0	1	1	0	1	1	1	0	0	1	1	1	1	1	0	0	0	0	1	0	0	0	1
ARL_HDCOG_TX17A	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ARL_ARPI_TX20	0	1	1	1	1	1	0	1	1	0	0	1	1	1	0	0	1	1	1	0	0	0	1
ARL_ARPI_TX22	0	1	1	1	1	1	0	1	1	0	0	1	1	1	0	0	1	1	1	0	0	0	1
ARL_ICB_CT2WS	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	1	1	0	0	1	1
ARL_ICB_RSVP	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	1	1
ARL_EEGCS_VEP	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	0	0	1	1
DCS_CANCTA_FT	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DCS_CANCTA_ODE	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0
NCTU_CANCTA_RWN_VDE	0	1	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0
TNO_CANCTA_ACC	1	0	1	1	0	1	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
TNO_CANCTA_FLERP	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Number of Studies : 27	2	25	10	14	6	8	2	13	10	3	7	11	17	12	7	2	7	10	18	1	1	5	7

1.4.1 Repeat subjects

The BCIT program required data collection at three different locations. Moreover, since each of the experiment teams at two of the locations were collecting data for 4 different experiments, some of the subjects enrolled multiple times. Thus, the subject numbers for the T2 and T3 experiments are "re-used" when the individual participates in more than one experiment within an overall study, in order to track the "same brain" when known. Information regarding repeat subjects was made available for the data processing team after it was de-identified. Table 8 and Table 9 indicate which subjects participated in multiple recording sessions, which means they were in the lab on multiple days. A value of "1" indicates valid data exists for the subject performing an experiment task.

Table 8. SANDR Repeat Subjects for BCIT Program Task 2

Subject ID	T2 X2 Experiment 1 session			T2 X6 Experiment 1 session			T2 X7 Experiment 1 session			T2 X8 Experiment 1 session				Total Runs	Total Sessions
	XC	XB	X2	XC	X6 CA	X6 CB	XC	X7 CA	X7 CB	XC	X8 CA	X8 CB	X8 CC		
2013	1	1	1				1	1	1					6	2
2015	1	1	1	1	1	1	1	1	1					9	3
2026	1		1				1	1	1					5	2
2029	1	1	1	1	1	1								6	2
2043				1	1	1	1	1	1	1	1	1	1	10	3
2044				1	1	1	1	1	1	1	1	1	1	10	3
2046				1	1	1	1	1	1					6	2
2048				1	1	1	1	1	1	1	1	1	1	10	3
2051				1	1	1	1	1	1	1	1	1	1	10	3
2055				1	1	1	1	1	1	1	1	1	1	10	3
2056				1	1	1	1	1	1	1	1	1	1	10	3
2061				1	1	1	1	1	1	1	1	1	1	10	3
2063				1	1	1	1	1	1	1	1	1	1	10	3

Table 9. SANDR Repeat Subjects for BCIT Program Task 3

Subject ID	T3 X1 Experiment 1 session			T3 X2 Experiment (longitudinal)												T3 X3 Experiment 1 session			T3 X4 Experiment 1 session			Total Runs	Total Sessions
	XC	XB	X1	XC (1 session)	X2	XC (1 session)	X2	XC (1 session)	X2	XC (1 session)	X2	XC (1 session)	X2	XC (1 session)	X2	XC	XB	X3	XC	XB	X4		
3103	1	1	1	1	1	1	1	1	1	1	1	1	1						1	1	1	16	7
3109	1	1	1																1	1	1	6	2
3115	1	1	1	1	1	1	1	1	1	1	1	1	1			1	1	1				16	7
3120	1	1	1	1	1	1	1	1	1	1	1	1	1							1	1	15	7
3201				1	1	1	1	1	1	1	1	1	1									10	5
3202				1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	18	8
3203				1	1	1	1	1	1	1	1	1	1			1	1	1	1	1	1	16	7
3204				1	1	1	1	1	1	1	1	1	1			1	1	1				13	6
3205				1	1	1	1	1	1	1	1	1	1									10	5
3206				1	1	1	1	1	1	1	1	1	1			1	1	1	1	1	1	16	7
3208				1	1	1	1	1	1	1	1	1	1						1	1	1	13	6
3210				1	1	1	1	1	1	1	1	1	1						1	1	1	13	6
3303																1	1	1	1	1	1	6	2
3306																1	1	1	1	1	1	6	2
3308																1	1	1	1	1	1	6	2
3311																1	1	1	1	1	1	6	2
3320	1	1	1													1	1	1	1	1	1	9	3
3402	1		1																1	1	1	5	2
3409	1	1	1																1	1	1	6	2
3412	1	1	1																1	1	1	6	2
3413	1	1	1																1	1	1	6	2
3422	1	1	1																1	1	1	6	2

For the RWNVDEDP program (Section 2.4.1.7), the 44 participants were paired into 22 driver/passenger dyads, with 40 subjects performing both driver and passenger tasks. Table 10 indicates which subject was the driver and which was the passenger for each of the 41 missions that produced useable data.

Table 10. Driver and Passenger Subjects for RWNVDEDP Experiments

Mission	Dyad	Driver	Passenger	File Date	File Time	Subject
1	1	3032	1021	20171213	81453	Driver
1	1	3032	1021	20171213	81453	Passenger
2	2	8019	7097	20171214	122546	Driver
2	2	8019	7097	20171214	122546	Passenger
3	3	5089	4086	20171219	82504	Driver
3	3	5089	4086	20171219	82504	Passenger
4	4	8114	8077	20171219	124011	Driver
4	4	8114	8077	20171219	124011	Passenger
5	5	2093	4069	20171220	123558	Driver
5	5	2093	4069	20171220	123558	Passenger
6	5	4069	2093	20180118	83907	Driver
6	5	4069	2093	20180118	83907	Passenger
7	2	7097	8019	20180118	125740	Driver
7	2	7097	8019	20180118	125740	Passenger
8	1	1021	3032	20180119	105948	Driver
8	1	1021	3032	20180119	105948	Passenger
9	3	4086	5089	20180129	125523	Driver
9	3	4086	5089	20180129	125523	Passenger
10	4	8077	8114	20180202	81945	Driver
10	4	8077	8114	20180202	81945	Passenger
11	7	8197	4275	20180207	150945	Driver
11	7	8197	4275	20180207	150945	Passenger
12	6	5189	4146	20180208	80343	Driver
12	6	5189	4146	20180208	80343	Passenger
13	8	8164	6128	20180209	102158	Driver
13	8	8164	6128	20180209	102158	Passenger
14	9	1139	4126	20180209	134025	Driver
14	9	1139	4126	20180209	134025	Passenger
15	10	2153	4207	20180214	123400	Driver
15	10	2153	4207	20180214	123400	Passenger
16	11	4251	1292	20180215	94227	Driver
16	11	4251	1292	20180215	94227	Passenger
17	12	5245	3284	20180220	94404	Driver
17	12	5245	3284	20180220	94404	Passenger
19	14	5226	2265	20180223	124236	Driver
19	14	5226	2265	20180223	124236	Passenger
20	8	6128	8164	20180301	90750	Driver
20	8	6128	8164	20180301	90750	Passenger
21	6	2093	4069	20180301	134021	Driver
21	6	2093	4069	20180301	134021	Passenger
22	15	2417	7427	20180307	91956	Driver
22	15	2417	7427	20180307	91956	Passenger
23	16	3303	3323	20180307	131548	Driver
23	16	3303	3323	20180307	131548	Passenger
24	18	5334	6351	20180309	85939	Driver
24	18	5334	6351	20180309	85939	Passenger
25	13	6172	8234	20180309	124700	Driver
25	13	6172	8234	20180309	124700	Passenger
26	22	5442	3333	20180314	84648	Driver
26	22	5442	3333	20180314	84648	Passenger
27	17	9407	7392	20180314	131617	Driver
27	17	9407	7392	20180314	131617	Passenger
28	11	1292	4251	20180315	94051	Driver
28	11	1292	4251	20180315	94051	Passenger
29	14	4126	1139	20180315	131539	Driver
29	14	4126	1139	20180315	131539	Passenger
30	20	3455	3437	20180319	91633	Driver
30	20	3455	3437	20180319	91633	Passenger
31	19	2468	2387	20180319	130018	Driver
31	19	2468	2387	20180319	130018	Passenger
32	21	8314	9346	20180322	130856	Driver
32	21	8314	9346	20180322	130856	Passenger

33	12	3284	5245	20180323	92327	Driver
33	12	3284	5245	20180323	92327	Passenger
34	15	7427	2417	20180323	131933	Driver
34	15	7427	2417	20180323	131933	Passenger
35	10	4207	2153	20180326	121215	Driver
35	10	4207	2153	20180326	121215	Passenger
36	7	4275	8197	20180326	152522	Driver
36	7	4275	8197	20180326	152522	Passenger
37	9	5089	4086	20180327	125823	Driver
37	9	5089	4086	20180327	125823	Passenger
38	16	1292	4251	20180328	130715	Driver
38	16	1292	4251	20180328	130715	Passenger
39	21	9346	8314	20180404	93712	Driver
39	21	9346	8314	20180404	93712	Passenger
41	22	3333	5442	20180409	93131	Driver
41	22	3333	5442	20180409	93131	Passenger
42	18	6351	5334	20180409	125638	Driver
42	18	6351	5334	20180409	125638	Passenger
43	19	2387	2468	20180410	125635	Driver
43	19	2387	2468	20180410	125635	Passenger

The ARPI program featured two studies, TX20 and TX22. Those data collection events featured some overlap of test participants, but it would require further investigation of the research induction paperwork in order to determine the corresponding subject ID number of the participants that were measured in both studies.

2 Study Descriptions

This section describes each dataset currently in the ARL SANDR, organized by research program.

2.1 Brain-Computer Interface Technology (BCIT)

The field of Brain-Computer Interaction Technologies may provide revolutionary Soldier-system capabilities. Much of the current research in these technologies tends to focus on developing communication and control technologies in the medical domain for assisting paralyzed or disabled individuals. These medical technologies augment function for these impaired clinical populations, but for healthy individuals these technologies provide inferior performance to using a manual interface. However, recent advancements show clear potential for developing technologies that support healthy populations and Soldier-system interaction by combining advancements in mobile brain signal technologies, signal processing methods and techniques, and low-power, lightweight computational capabilities.

The recent growth in this research area has led to the initiation of a Brain-Computer Interaction Technologies (BCIT) research /program as part of the U.S. Army Research Laboratory's Major Laboratory Program in Neuroscience, with the goal of developing and translating BCIT for Army-relevant applications. Within this context, a major focus area of the BCIT program is the development and validation of a fatigue-based performance prediction system for Army-relevant tasks.

2.1.1 BCIT Program Summary

This program provides for extended time-on-task measurements of subjects across various paradigms involving vigilance tasks, including driving (remaining and responding to vehicle perturbations), Rapid-Serial Visual Presentation (RSVP) target detection and acknowledgement based on images of designated targets and non-targets, and a guard duty assignment to control entry to a restricted area based on simulated identification information and requests for access.

The Calibration Driving and Baseline Driving tasks were performed by nearly all participants within the program, across three studies and laboratories, to generate sufficient data for large-scale analysis. The relatively short Calibration Driving runs collected at the beginning of the session can be contrasted with longer Baseline Driving runs collected later in the session, for each subject, as well as in the aggregate.

The three objectives of this program were to 1) develop a calibration and experimentation solution for fatigue-based performance prediction, 2) conduct experiments that will support ARL in the evaluation of the reliability and generalizability of a previously published fatigue-based driver performance prediction methodology (Lin, Wu, Jung, Liang, Huang, EURASIP Journal on Applied Signal Processing, 2005) in increasingly realistic driving tasks, 3) conduct experiments that will support ARL in the evaluation of the reliability and generalizability of this methodology to non-driving Army-relevant tasks. In support of these objectives, a set of three related experiments were specified, and conducted, at three different laboratories.

2.1.1.1 BCIT Program Task 1 (T1)

The Validation Phase experiment was designed to verify the driving simulator as a data collection platform for measurable driving behaviors using perturbation events and real-time calculations to measure lane deviation and heading deviation against the vehicle heading, and steering angle.

Perturbations (depicted in Figure 2-1) are a lateral force that increases in magnitude through a step function to “push” the vehicle to the left or the right until the subject responds, at which point the perturbation force scales down at a rate 3 times faster than it grows during onset. The subject response to a perturbation is known as a “driver correction” event, which begins when the subject turns into the direction of the force with a steering angle greater than 4 degrees.

Perturbation events have the following metrics of interest:

Reaction Time (RT):

Time from onset of perturbation (force > 0) to beginning of subject response

Response Time (or, Driver Correction Time):

Time from beginning to end of driver correction, where:

Driver correction begins when subject turns into perturbation force at steer angle > 4 degrees, and

Driver correction ends when (ABS(steering angle) < 1 degree && ABS(heading deviation) < 0.75 degrees)

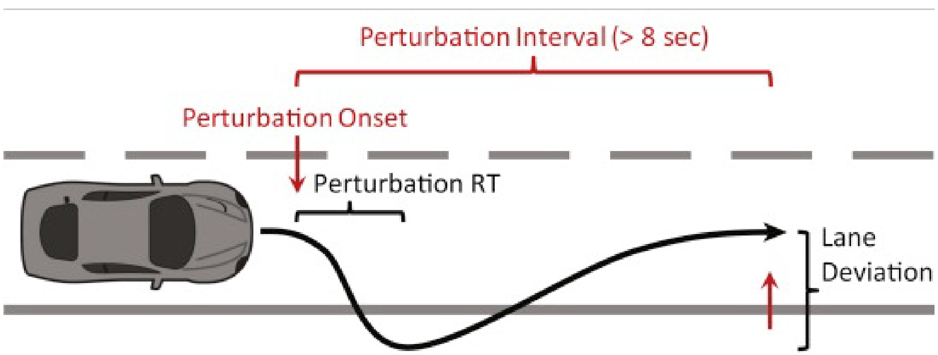


Figure 2-1. Vehicle Perturbation

2.1.1.1.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL/TARDEC
Protocol Number: ARL-20098-10051 (amended from TX18)
Protocol Name: “Army Relevant EEG Based Driver Performance Prediction”
Contract: W911NF-10-D-0002-0003

2.1.1.1.2 Location

This study was conducted at the ARL HRED laboratory located at Aberdeen Proving Ground, MD.

2.1.1.1.3 Subjects

A total of 25 subjects were recruited from the local government and contractor workforce.

2.1.1.1.4 Apparatus

Subjects performed driving tasks using a desktop driving simulator with steering wheel and foot pedals for vehicle control, (Real Time Technologies; Dearborn, MI), a video refresh rate of 900 Hz, and a vehicle state data sampling rate of 100 Hz (for log files).

They were instrumented with a BioSemi 64 (+8) EEG channel system with 4 eye and 2 mastoid channels recorded, with a sampling rate of 2048 Hz. Eye Tracking data was collected using a Sensomotoric Instruments (SMI) REDEYE250 system, with a 250 Hz sampling rate.

2.1.1.1.5 Demographics

Demographic information provided with this dataset includes subject ID, gender, birth year, age, dominant hand, height, and weight.

2.1.1.1.6 ARL BCIT T1 Experiment (Validation Phase) Tasks

The Validation Phase experiment was the only experiment within the study. For this experiment, subjects performed the following tasks during their session:

Calibration Driving (XC):
Duration: 15 minutes
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries.
Note: Speed controlled automatically by the driving simulator.

Baseline Driving (XB):
Duration: 60 minutes, or 45 minutes.
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries and maintain speed per speed limit signs along the road.

During this phase, the perturbation force was adjusted after the 8^{th} subject (of 25 total) to make them more noticeable. In addition, the duration of various driving “blocks” for the Baseline

Driving runs was explored. Subjects 1-8 drove 6 blocks of 10 minutes, with surveys administered between blocks. Subject 9 drove 2 blocks of 30 minutes with surveys administered between blocks, and subjects 10-26 (with 14 omitted) drove 1 block of 45 minutes, and were prompted verbally for subjective fatigue measures at the mid-point of the run.

2.1.1.1.7 POC

Brent Lance (brent.j.lance.civ@army.mil), ARL, Primary Investigator
Scott Kerick (scott.e.kerick.civ@army.mil), ARL, Associate Investigator
Justin Brooks (justin.r.brooks16.civ@army.mil), ARL, Data Analysis
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer
Mike Dunkel (mdunkel@omi.com), DCS Corp., Data Engineer

2.1.1.2 BCIT Program Task 2 (T2)

The study to Extend Lin et al. (2005) Approach to More Complex Driving Environments included a series of driving experiments, all of which featured perturbation events, but varied in other significant ways.

2.1.1.2.1 Protocol

The research protocol information for this study is as follows:

> Organization: ARL/Teledyne
> Protocol Number: ARL 12-040
> Protocol Name: “EEG-Based Measures of Driver Performance in Realistic Simulations"
> Contract: W911NF-10-D-0002-0003

2.1.1.2.2 Location

This study was conducted at the Teledyne laboratory located at their facility in Durham, NC.

2.1.1.2.3 Subjects

A total of 78 unique subjects were recruited from among the local population, via advertising. These subjects were recorded for 100 sessions, with some subjects participating in multiple experiments within the study.

2.1.1.2.4 Apparatus

Subjects performed driving tasks using a desktop driving simulator with steering wheel and foot pedals for vehicle control, (Real Time Technologies; Dearborn, MI), a video refresh rate of 900 Hz, and a vehicle state data sampling rate of 100 Hz (for log files).

They were instrumented with a BioSemi 64 (+8) EEG channel system with 4 eye and 2 mastoid channels recorded, with a sampling rate of 2048 Hz. Eye Tracking data was collected using a Sensomotoric Instruments (SMI) REDEYE250 system, with a 250 Hz sampling rate.

2.1.1.2.5 Demographics

Demographic information provided with this dataset includes subject ID, gender, birth year, age, dominant hand, height, and weight.

2.1.1.2.6 ARL BCIT T2 Experiments

Program Task 2 was comprised of four driving experiments, conducted using the same apparatus, but with variations in stimuli and expected subject behaviors.

2.1.1.2.6.1 Traffic Complexity (X2) Tasks

For this experiment, subjects performed the Calibration Driving task first, followed by the Baseline Driving task, and the Traffic Complexity task, with the sequence counter-balanced across subjects.

Calibration Driving (XC):
Duration: 15 minutes
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries.
Note: Speed controlled automatically by the driving simulator.

Baseline Driving (XB):
Duration: 45 minutes
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries and maintain speed per speed limit signs along the road.

Traffic Complexity (X2):
Duration: 45 minutes
Environment: Long, straight highway, visually complex environment, featuring oncoming traffic and traffic in the direction of travel (in the passing lane). Also had pedestrians on either side of the road, but not crossing the road. Figure 2-2 shows sample images of environmental entities contributing to the visual complexity.
Task requirements: Steer vehicle to maintain lane boundaries and maintain speed per speed limit signs along the road.

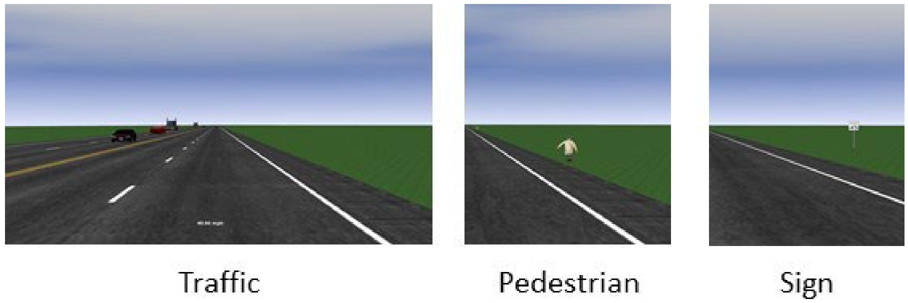


Figure 2-2. Traffic Complexity Environment

2.1.1.2.6.2 Speed Control (X6) Tasks

For this experiment, subjects performed the Calibration Driving task first, followed by two conditions of the Speed Control task, with the sequence counter-balanced across subjects.

> <u>Calibration Driving (XC)</u>:
> Duration: 15 minutes
> Environment: Long, straight highway in a visually sparse environment.
> Task requirements: Steer vehicle to maintain lane boundaries.
> Note: Speed controlled automatically by the driving simulator.
>
> <u>Speed Control (X6):</u>
> Duration: 45 minutes
> Environment: Long, straight highway in a visually sparse environment.
> Condition A Task requirements: Steer vehicle to maintain lane boundaries.
> Speed controlled automatically by the driving simulator (i.e., cruise control).
> Condition B Task requirements: Steer vehicle to maintain lane boundaries and maintain speed per speed limit signs along the road.

2.1.1.2.6.3 Auditory Cueing (X7) Tasks

For this experiment, subjects performed the Calibration Driving task first, followed by two conditions of the Auditory Cueing task, with the sequence counter-balanced across subjects.

> <u>Calibration Driving (XC):</u>
> Duration: 15 minutes
> Environment: Long, straight highway in a visually sparse environment.
> Task requirements: Steer vehicle to maintain lane boundaries.
> Note: Speed controlled automatically by the driving simulator.

Auditory Cueing (X7):
Duration: 45 minutes
Environment: Long, straight highway, visually sparse environment
Condition A and B Task requirements: Steer vehicle within lane boundaries and control speed per speed limit signs.
Condition A Stimulus: A short audio tone is played at random intervals throughout the run (not correlated with perturbation events). There's a check every 0.5 seconds to determine whether or not to play a cue. The chance of playing a cue is 5.625%. The probability of a cue happening is based on the average number of cues in the non-random condition, with the objective being to have a similar number of cues in both conditions.
Condition B Stimulus: A short audio tone is played prior to 90% of the scheduled perturbations, with the time delta between audio and perturbation onset varying randomly within a 2-second window.

2.1.1.2.6.4 Mind Wandering (X8) Tasks

For this experiment, subjects performed the Calibration Driving task first, followed by three conditions of the Mind Wandering task, with the sequence counter-balanced across subjects.

Calibration Driving (XC):
Duration: 15 minutes
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries.
Note: Speed controlled automatically by the driving simulator.

Mind Wandering (X8):
Duration: 30 minutes
Environment: Long, straight highway, visually complex environment with environmental traffic in both directions, including police cars, which represent visual targets.
Condition A, B, and C Task requirements: Steer vehicle within lane boundaries and control speed per speed limit signs. Press a button on the steering wheel when a police car is viewed.
Condition A Stimulus: Continuous audio is played during the run, which features task-related content (driving safety podcast) with external focus.
Condition B Stimulus: Continuous audio is played during the run, which features task-unrelated content (sport/news podcast) with external focus.
Condition C Stimulus: Continuous audio is played during the run, which features mindfulness content (meditation podcast) with internal focus.

2.1.1.2.7 POC

Scott Kerick (scott.e.kerick.civ@army.mil), ARL, Primary Investigator
Javier Garcia (javier.o.garcia.civ@army.mil), ARL, Data Analysis
Matthew Jaswa (mjaswa@dcscorp.com), System Developer
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer
Mike Dunkel (mdunkel@omi.com), DCS Corp., Data Engineer

2.1.1.3 BCIT Program Task 3 (T3)

The study to Extend Lin et al. (2005) Approach to Non-Driving Tasks included a series of four target detection experiments, featuring two different paradigms.

2.1.1.3.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL/SAIC
Protocol Number: ARL 12-041 (SAIC 2012040301)
Protocol Name: “Performance Prediction in Visual Tasks”
Contract: W911NF-10-D-0002-0003

2.1.1.3.2 Location

This study was conducted at the SAIC laboratory located at their facility in Louisville, CO.

2.1.1.3.3 Subjects

A total of 59 unique subjects were recruited from among the local population, via advertising. These subjects were recorded for 88 sessions. Some subjects participated in a longitudinal experiment, and some subjects participated in multiple experiments within the study.

2.1.1.3.4 Apparatus

Subjects performed driving tasks using a desktop driving simulator with steering wheel and foot pedals for vehicle control, (Real Time Technologies; Dearborn, MI), a video refresh rate of 900 Hz, and a vehicle state data sampling rate of 100 Hz (for log files). Subjects also performed RSVP and Guard Duty tasks using different, custom-designed software applications hosted on the same computer system.

They were instrumented with a BioSemi 256 (+8) EEG channel system with 4 eye and 2 mastoid channels recorded, with a sampling rate of 2048 Hz. Eye Tracking data was collected using a Sensomotoric Instruments (SMI) REDEYE250 system, with a 250 Hz sampling rate.

2.1.1.3.5 Demographics

Demographic information provided with this dataset includes subject ID, gender, birth year, age, dominant hand, height, and weight.

2.1.1.3.6 ARL BCIT T3 Experiments

Program Task 3 was comprised of two Rapid Serial Visual Presentation (RSVP) target detection experiment, and two guard duty experiments modeled from real-world tasks.

2.1.1.3.6.1 RSVP Baseline (X1) Tasks

RSVP Baseline study sessions include subjects performing the following tasks:

Calibration Driving (XC):
Duration: 15 minutes
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries.
Note: Speed controlled automatically by the driving simulator.

Baseline Driving (XB):
Duration: 60 minutes (6 10-minute blocks)
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries and maintain speed per speed limit signs along the road.

RSVP Baseline (X1):
Duration: 60 minutes (6 10-minute blocks)
Environment: Rapid Serial Visual Presentation of common objects as targets, including: chairs, containers, doors, posters, stairs
Task requirements: Press a button when a designated target is recognized
Stimulus requirements:
Different targets specified for each of blocks 1-5
The target for block 6 target = block 1 target
The 5 objects used in the images were chairs, containers, doors, posters, and stairs

Subjects performed the Calibration Driving task first, followed by the Baseline Driving task, and the RSVP Baseline task, with the sequence counter-balanced across subjects.

The Calibration Driving task and the Baseline Driving task are performed within this study to support the collection of data of a large number of subjects performing the same tasks for large-scale data analysis.

For the RSVP Baseline task, subjects were asked to view a series of images. The images were displayed in 6 blocks of 10 minutes, with a break in between block of approximately 2 minutes. For each block, one of the aforementioned objects was designated as the target, and images of the remaining objects were used as distractors. The subject was instructed to press a button each time a target image was perceived. Images contained chairs, containers, doors, posters, and stairs. One of these objects would be designated as the target for a given block, and the others would represent non-targets (for that block). The target was different for each block 1-5, and the target for block 6 was the same as it was for block 1. The sequence of target objects assigned to blocks 1-6 was counter-balanced across subjects. Figure 2-3 and Figure 2-4 contain samples from the image database.

Figure 2-3. Images of the Five Target Types for BCIT RVSP Experiments (X1 and X2)

Figure 2-4. Multiple Images of One Target Type for BCIT RVSP Experiments (X1 and X2)

The set of images used for RSVP stimuli were taken specifically for use in BCIT T3 experiments X1 and X2. Each picture has an associated set of attributes (see Figure 2-5) which specify information such as image luminance and target occlusion. There is a target viewer application, provided with the image database, which makes the attributes accessible via lookup. They can also be extracted programmatically, via the MATLAB Image Processing Toolbox (or possibly other, similar tools).

Parameter/DB Field	Target	NonTarget
Image filename	1xxxx jpg	2xxxx jpg
Image ID	1xxxx	2xxxx
Target mask file	1xxxx copy z.jpg	*
Indoor/outdoor	assigned	assigned
Clutter	1-5 assigned	1-5 assigned
Image luminance	0.299*µR+0.587*µG+0.114*µB	0.299*µR+0.587*µG+0.114*µB
# targets	computed	0
# target classes	computed	0
Target class	assigned	*
Target orientation	0°-180° assigned	*
Target occlusion	1-5 assigned	*
Camera distance	1-5 assigned	1-5 assigned
Target contrast	1-3 assigned	*
Target size	#target pixels /#image pixels (computed)	*
Centroid distance	distance from centroid to center of image (computed)	*

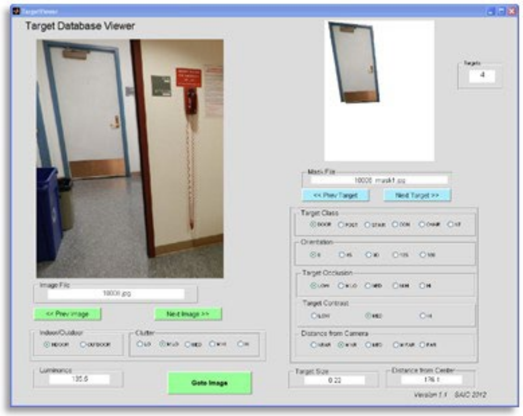
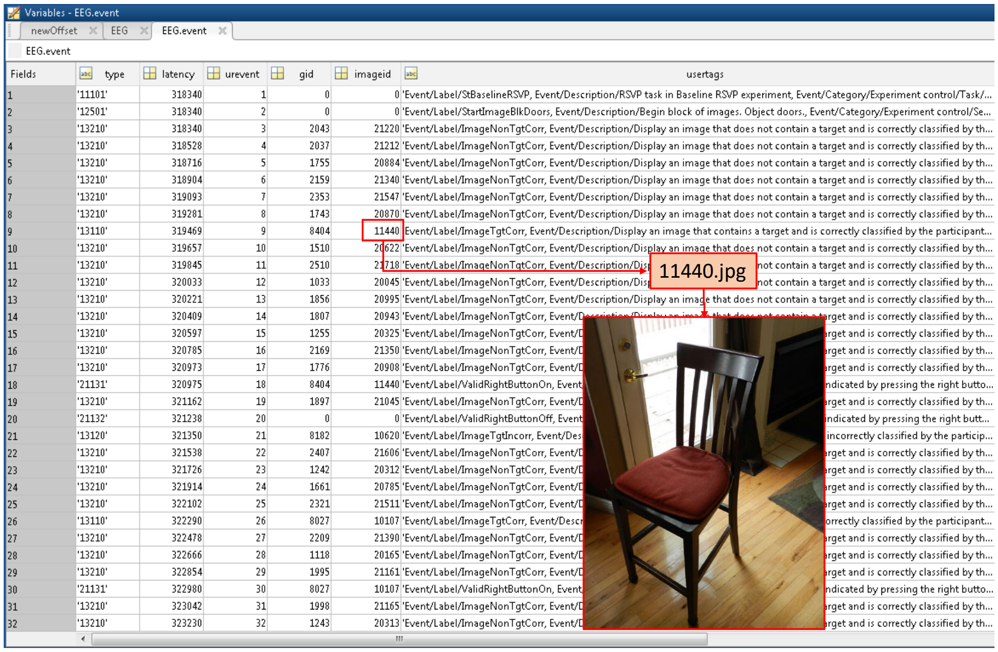


Figure 2-5. Target Image Attributes for BCIT RVSP Experiments (X1 and X2)

The stimuli are stored as jpg files with the image identifier as the name. This information is part of the event-related data stored in the EEG.event structure within the EEG file (see Figure 2-6).

Also stored in EEG.event is a 5-digit event code, in the column labeled 'type'. These codes are defined in the data specification spreadsheet, "BCI Data Specification v*nn*.xls", where represents the version number. The event code can be interpreted by reading from the event table, from left to right. For example, referencing Figure 2-7, it can be determined that an event code of 13110 (as in line 9 in Figure 2-6) represents the following:

Scenario | Present Image | Target | Correct classification (the last digit is not used in this case).



Variables - EEG.event

newOffset | EEG | EEG.event

EEG.event

Fields	type	latency	urevent	gid	imageid	usertags
1	'11101'	318340	1	0	0	'Event/Label/StBaselineRSVP, Event/Description/RSVP task in Baseline RSVP experiment, Event/Category/Experiment control/Task/...
2	'12501'	318340	2	0	0	'Event/Label/StartImageBlkDoors, Event/Description/Begin block of images. Object doors., Event/Category/Experiment control/Se...
3	'13210'	318340	3	2043	21220	'Event/Label/ImageNonTgtCorr, Event/Description/Display an image that does not contain a target and is correctly classified by th...
4	'13210'	318528	4	2037	21212	'Event/Label/ImageNonTgtCorr, Event/Description/Display an image that does not contain a target and is correctly classified by th...
5	'13210'	318716	5	1755	20884	'Event/Label/ImageNonTgtCorr, Event/Description/Display an image that does not contain a target and is correctly classified by th...
6	'13210'	318904	6	2159	21340	'Event/Label/ImageNonTgtCorr, Event/Description/Display an image that does not contain a target and is correctly classified by th...
7	'13210'	319093	7	2353	21547	'Event/Label/ImageNonTgtCorr, Event/Description/Display an image that does not contain a target and is correctly classified by th...
8	'13210'	319281	8	1743	20870	'Event/Label/ImageNonTgtCorr, Event/Description/Display an image that does not contain a target and is correctly classified by th...
9	'13110'	319469	9	8404	11440	Event/Label/ImageTgtCorr, Event/Description/Display an image that contains a target and is correctly classified by the participant...
10	'13210'	319657	10	1510	20622	'Event/Label/ImageNonTgtCorr, Event/Description/Display an image that does not contain a target and is correctly classified by th...
11	'13210'	319845	11	2510	21718	'Event/Label/ImageNonTgtCorr, Event/Description/Disp … not contain a target and is correctly classified by th...
12	'13210'	320033	12	1033	20045	'Event/Label/ImageNonTgtCorr, Event/Description/Dis … not contain a target and is correctly classified by th...
13	'13210'	320221	13	1856	20995	'Event/Label/ImageNonTgtCorr, Event/Description/Display an image that does not contain a target and is correctly classified by th...
14	'13210'	320409	14	1807	20943	'Event/Label/ImageNonTgtCorr, Event/ … rget and is correctly classified by th...
15	'13210'	320597	15	1255	20325	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
16	'13210'	320785	16	2169	21350	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
17	'13210'	320973	17	1776	20908	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
18	'21131'	320975	18	8404	11440	'Event/Label/ValidRightButtonOn, Event … ndicated by pressing the right butto...
19	'13210'	321162	19	1897	21045	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
20	'21132'	321238	20	0	0	'Event/Label/ValidRightButtonOff, Event … ndicated by pressing the right butt...
21	'13120'	321350	21	8182	10620	'Event/Label/ImageTgtIncorr, Event/Des … incorrectly classified by the particip...
22	'13210'	321538	22	2407	21606	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
23	'13210'	321726	23	1242	20312	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
24	'13210'	321914	24	1661	20785	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
25	'13210'	322102	25	2321	21511	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
26	'13110'	322290	26	8027	10107	'Event/Label/ImageTgtCorr, Event/Descr … orrectly classified by the participant...
27	'13210'	322478	27	2209	21390	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
28	'13210'	322666	28	1118	20165	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
29	'13210'	322854	29	1995	21161	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
30	'21131'	322980	30	8027	10107	'Event/Label/ValidRightButtonOn, Event … ndicated by pressing the right butto...
31	'13210'	323042	31	1998	21165	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...
32	'13210'	323230	32	1243	20313	'Event/Label/ImageNonTgtCorr, Event/D … rget and is correctly classified by th...

Figure 2-6. BCIT RVSP (X1 and X2) EEG.event Data

EventType	EventSubtype	Variant	EventSubvariant	Demarcation
1 Scenario	n Event Subtype	n ScenarioType		n Demarcation
	1 Execute Scenario	1 Baseline	0 not used	1 Onset
		2 Expertise		2 Offset
		n TargetObject		n Demarcation
	2 Present Block of Images	1 Object Stairs	0 not used	1 Onset
		2 Object Containers		2 Offset
		3 Object Posters		
		4 Object Chairs		
		5 Object Doors		
		n ImageContent	n ImageClassification	0 not used
	3 Present Image	1 Target	1 Correct Classification	*instantaneous*
		2 Non-target	2 Incorrect Classification	
			3 Indeterminate Classification	
2 Behavioral	n Event Subtype	n ButtonHand	n ButtonPushResult	n Demarcation
	1 Button Push	1 Right Button	1 Indeterminate	1 Onset
		2 Left Button	2 False Alarm	2 Offset
			3 Valid Detection	
			4 Repeated Detection	
3 External	n Event Subtype	n InstructionType		n Demarcation
	1 Instruction	1 Take Survey	0 not used	1 Onset
				2 Offset
		n SystemFailureType		n Demarcation
	2 System Error	0 Other	0 not used	1 Onset
		1 Experimental System Failed		2 Offset
		2 Arduino Failed		
		3 EEG System Failed		
		4 Eye Tracking System Failed		

Figure 2-7. BCIT RVSP Event Codes

The indicator "correct classification" is based on post-processing results that determined whether the subject correctly identified the image as a target.

2.1.1.3.6.2 RSVP Expertise (X2) Tasks

RSVP Expertise study sessions include subjects performing the following tasks:

Calibration Driving (XC) (days 1-5)
Duration: 15 minutes
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries.
Note: Speed controlled automatically by the driving simulator.

Baseline Driving (XB) (day 1 only):
Duration: 60 minutes (6 10-minute blocks)
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries and maintain speed per speed limit signs along the road.

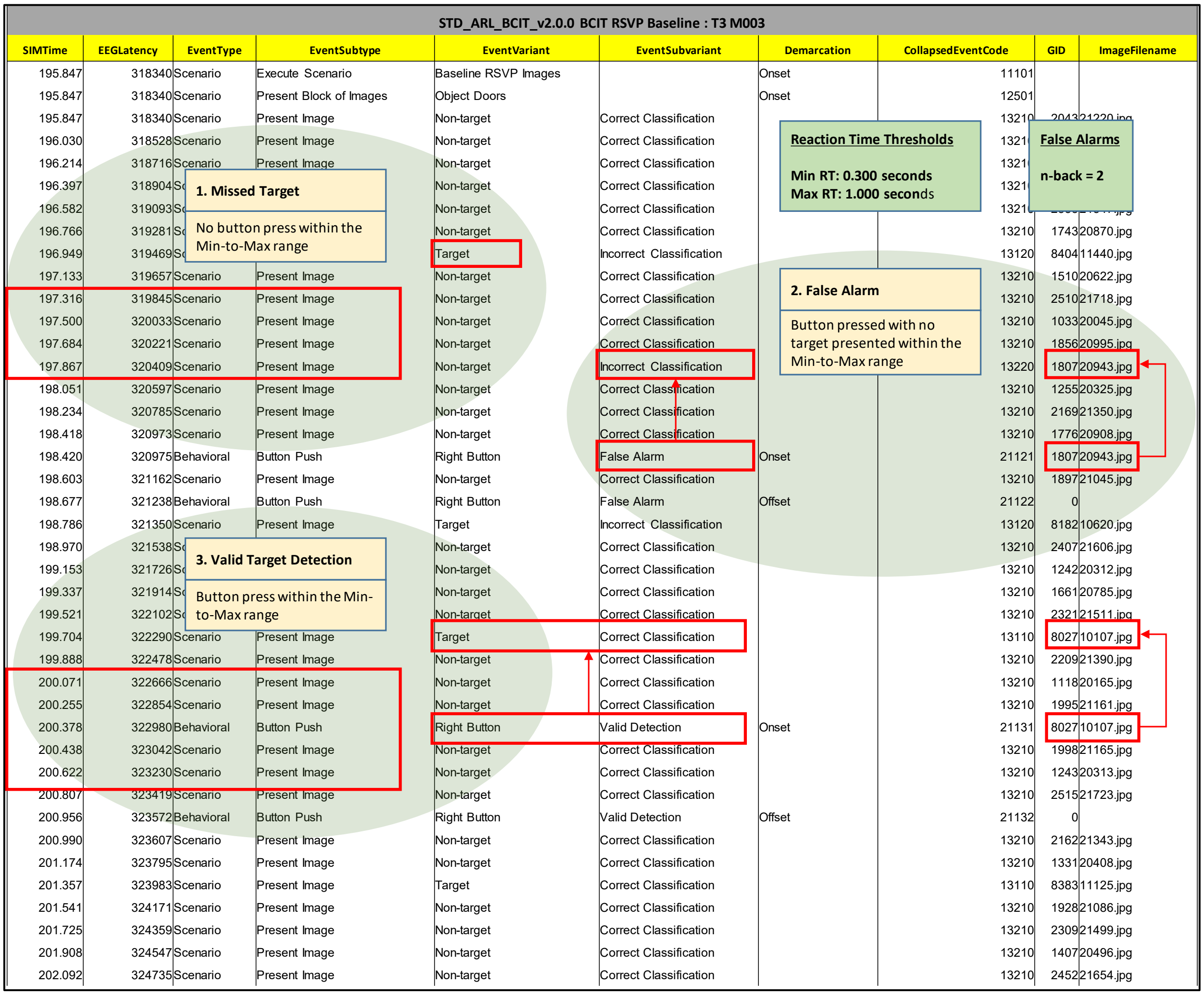
RSVP Expertise (X2) (days 1-5):
Duration: 60 minutes (6 10-minute blocks)
Environment: Rapid Serial Visual Presentation of common objects as targets, including: chairs, containers, doors, posters, stairs
Task requirements: Press a button when a designated target is recognized
Stimulus requirements:

> The target was different for each day of data collection (1-5), and within each day, the target was the same for all blocks (1-6)
> Designated target objects for a day were chairs, containers, doors, posters, and stairs
> The sequence of target-to-day assignments was counter-balanced across subjects

Subjects performed the Calibration Driving task first, followed by the Baseline Driving task, and then the RSVP Expertise task on day 1 of the 5-day study. On days 2-5, subjects performed the Calibration Driving task followed by the RSVP Expertise task.

The Calibration Driving task and the Baseline Driving task are performed within this study to support the collection of data of a large number of subjects performing the same tasks for large-scale data analysis.

For the RSVP Expertise task, subjects were asked to view a rapid serial visual presentation (RSVP) of images, all of which contained one or more of the following objects: chairs, containers, doors, posters, or stairs. The same image database used for the stimuli in the RSVP Baseline experiment (X1) was used for this experiment. The images were displayed in 6 blocks of 10 minutes, with a break in between block of approximately 2 minutes. One of the aforementioned objects was designated as the target for all six blocks, and images of the remaining objects were used as distractors. The subject was instructed to press a button each time a target image was perceived. For a subject's four remaining data collection sessions, each individual would return on a different day at roughly the same time of day for each session and begin with the Calibration Driving task. For each of the 5 RSVP Expertise recording sessions, a different target was designated, and used for all 6 image presentation blocks in the task.



STD_ARL_BCIT_v2.0.0 BCIT RSVP Baseline : T3 M003

SIMTime	EEGLatency	EventType	EventSubtype	EventVariant	EventSubvariant	Demarcation	CollapsedEventCode	GID	ImageFilename
195.847	318340	Scenario	Execute Scenario	Baseline RSVP Images		Onset	11101		
195.847	318340	Scenario	Present Block of Images	Object Doors		Onset	12501		
195.847	318340	Scenario	Present Image	Non-target	Correct Classification		13210	2043	21220.jpg
196.030	318528	Scenario	Present Image	Non-target	Correct Classification		1321[illegible]	[illegible]	[illegible]
196.214	318716	Scenario	Present Image	Non-target	Correct Classification		1321[illegible]	[illegible]	[illegible]
196.397	318904	Sc[illegible]	[illegible]	Non-target	Correct Classification		1321[illegible]	[illegible]	[illegible]
196.582	319093	Sc[illegible]	[illegible]	Non-target	Correct Classification		1321[illegible]	[illegible]	[illegible]
196.766	319281	Sc[illegible]	[illegible]	Non-target	Correct Classification		13210	1743	20870.jpg
196.949	319469	Sc[illegible]	[illegible]	Target	Incorrect Classification		13120	8404	11440.jpg
197.133	319657	Scenario	Present Image	Non-target	Correct Classification		13210	1510	20622.jpg
197.316	319845	Scenario	Present Image	Non-target	Correct Classification		13210	2510	21718.jpg
197.500	320033	Scenario	Present Image	Non-target	Correct Classification		13210	1033	20045.jpg
197.684	320221	Scenario	Present Image	Non-target	Correct Classification		13210	1856	20995.jpg
197.867	320409	Scenario	Present Image	Non-target	Incorrect Classification		13220	1807	20943.jpg
198.051	320597	Scenario	Present Image	Non-target	Correct Classification		13210	1255	20325.jpg
198.234	320785	Scenario	Present Image	Non-target	Correct Classification		13210	2169	21350.jpg
198.418	320973	Scenario	Present Image	Non-target	Correct Classification		13210	1776	20908.jpg
198.420	320975	Behavioral	Button Push	Right Button	False Alarm	Onset	21121	1807	20943.jpg
198.603	321162	Scenario	Present Image	Non-target	Correct Classification		13210	1897	21045.jpg
198.677	321238	Behavioral	Button Push	Right Button	False Alarm	Offset	21122	0	
198.786	321350	Scenario	Present Image	Target	Incorrect Classification		13120	8182	10620.jpg
198.970	321538	Sc[illegible]	[illegible]	Non-target	Correct Classification		13210	2407	21606.jpg
199.153	321726	Sc[illegible]	[illegible]	Non-target	Correct Classification		13210	1242	20312.jpg
199.337	321914	Sc[illegible]	[illegible]	Non-target	Correct Classification		13210	1661	20785.jpg
199.521	322102	Sc[illegible]	[illegible]	Non-target	Correct Classification		13210	2321	21511.jpg
199.704	322290	Scenario	Present Image	Target	Correct Classification		13110	8027	10107.jpg
199.888	322478	Scenario	Present Image	Non-target	Correct Classification		13210	2209	21390.jpg
200.071	322666	Scenario	Present Image	Non-target	Correct Classification		13210	1118	20165.jpg
200.255	322854	Scenario	Present Image	Non-target	Correct Classification		13210	1995	21161.jpg
200.378	322980	Behavioral	Button Push	Right Button	Valid Detection	Onset	21131	8027	10107.jpg
200.438	323042	Scenario	Present Image	Non-target	Correct Classification		13210	1998	21165.jpg
200.622	323230	Scenario	Present Image	Non-target	Correct Classification		13210	1243	20313.jpg
200.807	323419	Scenario	Present Image	Non-target	Correct Classification		13210	2515	21723.jpg
200.956	323572	Behavioral	Button Push	Right Button	Valid Detection	Offset	21132	0	
200.990	323607	Scenario	Present Image	Non-target	Correct Classification		13210	2162	21343.jpg
201.174	323795	Scenario	Present Image	Non-target	Correct Classification		13210	1331	20408.jpg
201.357	323983	Scenario	Present Image	Target	Correct Classification		13110	8383	11125.jpg
201.541	324171	Scenario	Present Image	Non-target	Correct Classification		13210	1928	21086.jpg
201.725	324359	Scenario	Present Image	Non-target	Correct Classification		13210	2309	21499.jpg
201.908	324547	Scenario	Present Image	Non-target	Correct Classification		13210	1407	20496.jpg
202.092	324735	Scenario	Present Image	Non-target	Correct Classification		13210	2452	21654.jpg

Figure 2-8. BCIT RVSP Stimuli-Response Correlation

The relationship between stimuli and subject response was ascertained during post-processing using a Minimum Reaction Time threshold of 0.3 seconds, and a Maximum Reaction Time threshold of 1.0 seconds. These threshold values create a “time window”, when added to the time of the target image presentation that the subject response (button press) must occur within in order to be considered a valid detection. Target images that have a corresponding button press within the reaction time range are considered valid detections, and the target stimulus is associated with the button press event via the image id. Target images that have no corresponding button press within the reaction time range are considered invalid detections (aka missed targets). When a button press occurs and there was no corresponding target image presented within the reaction time range from the button press, the response is also deemed an invalid detection (aka false alarm). In this case, the subject response is associated with a non-target image using an n-back scheme, where n is set to a value of two. The image associated with the false alarm is estimated by subtracting the n-back value from the length of the list of non-target images within the reaction time range from button press. Each of the three event types described is depicted in Figure 2-8.

2.1.1.3.6.3 Basic Guard Duty (X3) Tasks

Basic Guard Duty study sessions include subjects performing the following tasks:

Calibration Driving (XC):
Duration: 15 minutes
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries.
Note: Speed controlled automatically by the driving simulator.

Baseline Driving (XB):
Duration: 60 minutes (6 10-minute blocks)
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries and maintain speed per speed limit signs along the road.

Basic Guard Duty (X3):
Duration: 50 minutes (10 5-minute blocks)
Environment: A representative identification card with a face picture and other information is displayed alongside another face image.
Task requirements: Press a button to allow or deny access to a restricted area to the person pictured in the screen presentation, possessing the accompanying ID card.

Subjects performed the Calibration Driving task first, followed by the Baseline Driving task, and the Basic Guard Duty task, with the sequence counter-balanced across subjects.

The Calibration Driving task and the Baseline Driving task are performed within this study to support the collection of data of a large number of subjects performing the same tasks for large-scale data analysis.

The Basic Guard Duty task entailed a serial presentation of replica identification (ID) cards (750 × 450 pixels) paired with a reference image (300 × 400 pixels). The replica ID cards had eight components or fields in addition to a common background. These components were photo, name, date of birth (DOB), date of issue, date of expiration, area access, ID number, bar code and watermark. The reference images consisted of color photographs of faces. Both the ID photo and reference image were chosen from the Multi-PIE database (Gross, Matthews, Cohn, Kanade, & Baker, 2010). This database consists of color photographs (forward facing headshots) of individuals taken at different points in time. Therefore, while the ID photo and reference image were of the same individual, the images were not identical (e.g., different hairstyle, different clothes, different lighting). Figure 2-9 depicts a sample of the task stimulus.

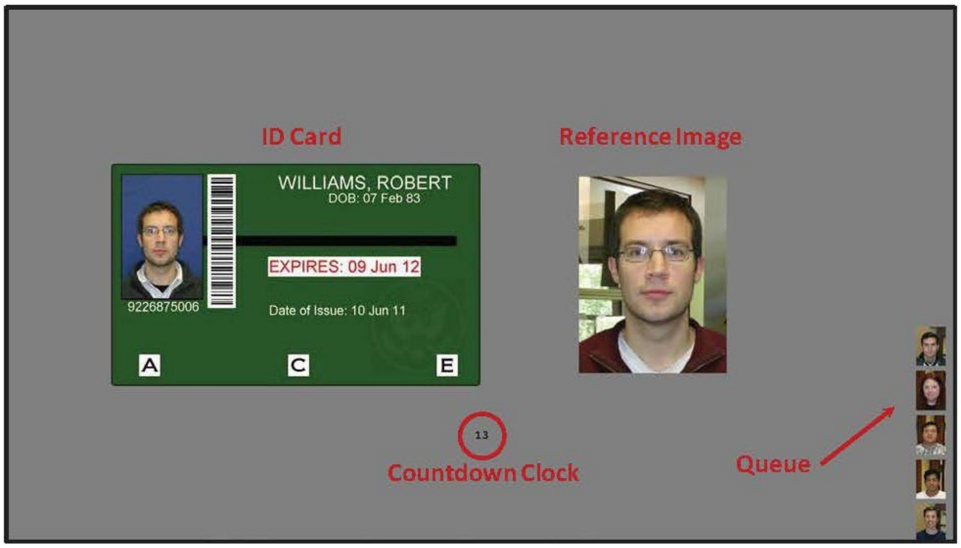


Figure 2-9. Screen Shot of Guard Duty Task (ID Card and Request for Access)

The task was divided into ten blocks of five minutes each. At the beginning of each block, participants were instructed that they were guarding a restricted area that required a particular letter designation on the ID card for access (e.g., area C access required). Participants were asked to determine if the individual in the image, paired with the corresponding ID card, should have access to their restricted area. Some of the ID cards were valid and some were not (e.g., expiration date passed, incorrect access area, or photos did not match). Participants were instructed to press either an "allow" or "deny" button for each image-ID pairing. The two-alternative forced-choice response was self-paced with a maximum time limit of 20 s. If the participants chose to deny access, they were subsequently asked to provide a reason. Reasons for denied access were selected from a numerical list of five options: 1—incorrect access, 2—expired ID, 3—suspicious DOB, 4—face mismatch, 5—no watermark. If the participant did not respond within the allotted time, the computer forced a "deny" decision. For the Basic Guard Duty task, the restricted area (area A–E) assigned at the beginning of the first block of images remained unchanged through all ten blocks. To maintain consistency across participants, expiration dates were automatically generated at the beginning of the experiment to have a symmetrical distribution around the current date. This distribution was such that the majority of IDs had expiration dates temporally close to the current date (i.e., in the near future or recent past).

In each block, the image-ID pairings were presented at one of six different stochastic queuing rates, ranging from 1 to 25 per minute (1, 2.5, 10, 15, 20, and 25 per minute). The queuing rate varied within each block according to a predefined profile. The rate profile had randomly permuted epochs of each queuing rate. Each epoch lasted 30 s with approximately twice as many low rate epochs (1 and 2.5 image-IDs per minute) as high. The rate profiles were shifted for each participant (Latin square design) so that each rate profile was assigned to every block for at least two participants. The current rate was indicated through a processing queue, on the extreme right-hand side of the display, notifying each participant how many IDs are waiting to be checked. For slow rates, most participants were able to process all IDs in their queue and had periods where they were waiting for the next ID (i.e., blank screen). For fast rates, most participants were not able to processes IDs as quickly as they were added to the queue, increasing the size of the processing queue. IDs in the queue persisted until they were processed by the participant or the block ended. At the beginning of the experiment, participants were instructed to correctly process each image-ID while keeping the queue as short as possible. Whereas the stochastic queuing rate was used to increase task realism, incorporating periods of high and low task demand, the dynamic rate itself was not explicitly considered an independent factor in the present study.

All blocks contained the same ratio of valid and invalid image-ID pairings (82% valid, 18% invalid). The majority of invalid IDs were due to incorrect access (6%) and expiration (6%) whereas the rest were invalid for the other reasons: suspicious DOB (2%), face mismatch (2%), no watermark (2%). This second group of invalid IDs served as catch trials to verify that participants were examining all fields of the ID.

2.1.1.3.6.4 Advanced Guard Duty (X3) Tasks

Advanced Guard Duty study sessions include subjects performing the following tasks:

Calibration Driving (XC):
Duration: 15 minutes
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries.
Note: Speed controlled automatically by the driving simulator.

Baseline Driving (XB):
Duration: 60 minutes (6 10-minute blocks)
Environment: Long, straight highway in a visually sparse environment.
Task requirements: Steer vehicle to maintain lane boundaries and maintain speed per speed limit signs along the road.

Advanced Guard Duty (X4):
Duration: 50 minutes (10 5-minute blocks)
Environment: A representative identification card with a face picture and other information is displayed alongside another face image.
Task requirements: Press a button to allow or deny access to a restricted area to the person pictured in the screen presentation, possessing the accompanying ID card.

Subjects performed the Calibration Driving task first, followed by the Baseline Driving task, and the Advanced Guard Duty task, with the sequence counter-balanced across subjects.

The Calibration Driving task and the Baseline Driving task are performed within this study to support the collection of data of a large number of subjects performing the same tasks for large-scale data analysis.

The guard duty task entailed a serial presentation of replica identification (ID) cards (750 × 450 pixels) paired with a reference image (300 × 400 pixels). The replica ID cards had eight components or fields in addition to a common background. These components were photo, name, date of birth (DOB), date of issue, date of expiration, area access, ID number, bar code and watermark. The reference images consisted of color photographs of faces. Both the ID photo and reference image were chosen from the Multi-PIE database (Gross, Matthews, Cohn, Kanade, & Baker, 2010). This database consists of color photographs (forward facing headshots) of individuals taken at different points in time. Therefore, while the ID photo and reference image were of the same individual, the images were not identical (e.g., different hairstyle, different clothes, different lighting). The task was divided into ten blocks of five minutes each. The same ID and faces database used for the stimuli in the Basic Guard Duty experiment (X3) was used for this experiment.

At the beginning of each block, participants were instructed that they were guarding a restricted area that required a particular letter designation on the ID card for access (e.g., area C access required). Participants were asked to determine if the individual in the image, paired with the corresponding ID card, should have access to their restricted area. Some of the ID cards were valid and some were not (e.g., expiration date passed, incorrect access area, or photos did not match). Participants were instructed to press either an "allow" or "deny" button for each image-ID pairing. The two-alternative forced-choice response was self-paced with a maximum time limit of 20 s. If the participants chose to deny access, they were subsequently asked to provide a reason. Reasons for denied access were selected from a numerical list of five options: 1—incorrect access, 2—expired ID, 3—suspicious DOB, 4—face mismatch, 5—no watermark. If the participant did not respond within the allotted time, the computer forced a "deny" decision. The restricted area (area A–E) assigned at the beginning of each block was randomly chosen without replacement such that all participants completed two blocks guarding each of the five areas. To maintain consistency across participants, expiration dates were automatically generated at the beginning of the experiment to have a symmetrical distribution around the current date. This distribution was such that the majority of IDs had expiration dates temporally close to the current date (i.e., in the near future or recent past).

In each block, the image-ID pairings were presented at one of six different stochastic queuing rates, ranging from 1 to 25 per minute (1, 2.5, 10, 15, 20, and 25 per minute). The queuing rate varied within each block according to a predefined profile. The rate profile had randomly permuted epochs of each queuing rate. Each epoch lasted 30 s with approximately twice as many low rate epochs (1 and 2.5 image-IDs per minute) as high. The rate profiles were shifted for each participant (Latin square design) so that each rate profile was assigned to every block for at least two participants. The current rate was indicated through a processing queue, on the extreme right-hand side of the display, notifying each participant how many IDs are waiting to be checked. For slow rates, most participants were able to process all IDs in their queue and had periods where they were waiting for the next ID (i.e., blank screen). For fast rates, most participants were not able to processes IDs as quickly as they were added to the queue, increasing the size of the processing queue. IDs in the queue persisted until they were processed by the participant or the block ended. At the beginning of the experiment, participants were instructed to correctly process each image-ID while keeping the queue as short as possible. Whereas the stochastic queuing rate was used to increase task realism, incorporating periods of high and low task demand, the dynamic rate itself was not explicitly considered an independent factor in the present study.

All blocks contained the same ratio of valid and invalid image-ID pairings (82% valid, 18% invalid). The majority of invalid IDs were due to incorrect access (6%) and expiration (6%) whereas the rest were invalid for the other reasons: suspicious DOB (2%), face mismatch (2%), no watermark (2%). This second group of invalid IDs served as catch trials to verify that participants were examining all fields of the ID.

2.1.1.3.7 POC

Jon Touryan (jonathan.o.touryan.civ@army.mil), System Developer, Researcher, Analyst
Anthony Reis (anthony.j.ries2.civ@army.mil), ARL, Researcher
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer
Mike Dunkel (mdunkel@omi.com), DCS Corp., Data Engineer

2.1.2 BCIT Publications

Gregory Apker, Brent Lance, Scott Kerick, & Kaleb McDowell (2013) **Combined Linear Regression and Quadratic Classification Approach for an EEG-Based Prediction of Driver**

Performance. *In: Schmorrow D.D., Fidopiastis C.M. (eds) Foundations of Augmented Cognition. AC 2013. Lecture Notes in Computer Science, vol 8027*. Springer, Berlin, Heidelberg
DOI: 10.1007/978-3-642-39454-6_24
http://link.springer.com/chapter/10.1007/978-3-642-39454-6_24

Justin Brooks, Scott Kerick, & Kaleb McDowell (2013). **Event-related alpha desynchronization related to the scaling of steering wheel corrections**. *First International SomnoAlert Conference, Brussels Belgium, 22-24, Feb 2014*.
DOI: NA
http://www.somnosafe.com/sites/default/files/fichiers/somnosafe-somnoalertabstracts-20140221.pdf

Vernon Lawhern, Scott Kerick, & Kay Robbins (2013) **Detecting alpha spindle events in EEG time series using adaptive autoregressive models**. *BMC Neuroscience 2013* 14:101
DOI: 10.1186/1471-2202-14-101
http://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-14-101

Jonathan Touryan, Gregory Apker, Scott Kerick, & Kaleb McDowell (2013) **Translation of EEG-Based Performance Prediction Models to Rapid Serial Visual Presentation Tasks**. *Foundations of Augmented Cognition: Lecture Notes in Computer Science Volume 8027, Chapter: Translation of EEG-based performance prediction models to rapid serial visual presentation tasks*, Publisher: Springer Berlin Heidelberg, pp.521-530
DOI: 10.1007/978-3-642-39454-6_56
https://www.researchgate.net/publication/249658359_Translation_of_EEG-Based_Performance_Prediction_Models_to_Rapid_Serial_Visual_Presentation_Tasks

Gregory Apker, Brent Lance, Scott Kerick, & Kaleb McDowell (2014) **Estimating Driver Performance Using Multiple Electroencephalography (EEG)-Based Regression Algorithms**. *U.S. Army Research Laboratory Technical Report Series, ARL-TR-7074*
DOI: NA
www.dtic.mil/get-tr-doc/pdf?AD=ADA609346

J. Touryan, G. Apker, B.J. Lance, S.E. Kerick, A.J. Ries, & K. McDowell (2014) **Common EEG features for behavioral estimation in disparate, real-world tasks**. *Biological Psychology*
DOI: 10.1016/j.biopsycho.2015.12.009
https://doi.org/10.1016/j.biopsycho.2015.12.009

J. Touryan, G. Apker, B.J. Lance, S.E. Kerick, A.J. Ries, & K. McDowell (2014) **Estimating endogenous changes in task performance from EEG**. *Frontiers in Neuroscience | Neuroprosthetics*
DOI: 10.3389/fnins.2014.00155
http://dx.doi.org/10.3389/fnins.2014.00155

N. Bigdely-Shamlo, T. Mullen, C. Kothe, K.-M. Su, & K. Robbins (2015). **The PREP pipeline: standardized preprocessing for large-scale EEG analysis**, *Frontiers in Neuroinformatics 9*
DOI: 10.3389/fninf.2015.00016
http://dx.doi.org/10.3389/fninf.2015.00016

Justin R. Brooks, Scott E. Kerick, & Kaleb McDowell (2015) **Novel Measure of Driver and Vehicle Interaction Demonstrates Transient Changes Related to Alerting**, *Journal of Motor Behavior*, 47:2, 106-116
DOI: 10.1080/00222895.2014.959887
http://www.tandfonline.com/doi/abs/10.1080/00222895.2014.959887

Justin Brooks, Scott Kerick (2015) **Event-related alpha perturbations related to the scaling of steering wheel corrections**, *Physiology & Behavior 149* (pp287-293)
DOI: 10.1016/j.physbeh.2015.05.026
https://doi.org/10.1016/j.physbeh.2015.05.026

J.O. Garcia, J. Brooks, S. Kerick, T. Johnson, T.R. Mullen, and J.M. Vettel (2015) **Estimating direction in bra in-behavior interactions: Proactive and reactive brain states in driving**. *NeuroImage 150* (pp239-249)
DOI: 10.1016/j.neuroimage.2017.02.057
https://doi.org/10.1016/j.neuroimage.2017.02.057

K. Kleifges, N. Bigdely-Shamlo, S. Kerick, & K. Robbins (2017). **BLINKER: Automated extraction of ocular indices from EEG enabling large-scale analysis**. *Frontiers in Neuroscience: Neurotechnology*. 03 February 2017
DOI: 10.3389/fnins.2017.00012
https://doi.org/10.3389/fnins.2017.00012.

2.1.3 BCIT Datasets

A total of 701 datasets were collected across 213 recording sessions, using 162 unique subjects from three different laboratories (ARL, Teledyne, SAIC) within the BCIT study. A small portion of the data had unrecoverable errors, leaving 688 datasets from 210 recording sessions, with 159 unique subjects. Of those subjects, 34 participated in multiple sessions, due to either the longitudinal study, signing up for multiple experiments, or both. The subject ID associated with each dataset is unique within the BCIT study.

There are 662 complete datasets, 18 that were interrupted by a system failure, but resumed, and 8 that were abbreviated runs. The total size of the EEG files is 2,661.36 GB, representing 453.21 hours of EEG recording.

2.1.3.1 ARL_BCIT_CalibrationDriving Dataset

A total of 247 datasets were collected across 247 recording sessions. This data was collected from 156 unique subjects, and came from three different laboratories (ARL, Teledyne, SAIC) within the BCIT study. The subject ID can be used to identify data from subjects that participated in the RSVP Expertise longitudinal study, and those who participated in multiple experiments within one of the program tasks.

All datasets are complete. None were abbreviated or interrupted. The total size of the EEG files is 380.80 GB, representing 63.56 hours of EEG recording.

2.1.3.2 ARL_BCIT_BaselineDriving Dataset

A total of 128 datasets were collected across 128 recording sessions. This data was collected from 109 unique subjects, and came from three different laboratories (ARL, Teledyne, SAIC) within the BCIT study. The subject ID can be used to identify data from subjects that participated in multiple experiments within one of the program tasks.

There are 123 complete datasets, 3 that were interrupted by a system failure, but resumed, and 2 that were abbreviated runs. The total size of the EEG files is 883.40 GB, representing 131.47 hours of EEG recording.

2.1.3.3 ARL_BCIT_TrafficComplexity Dataset

A total of 30 datasets were collected across 30 recording sessions, from 30 unique subjects recruited via the Teledyne laboratory. The subject ID can be used to identify data from subjects that participated in multiple experiments within program task 2.

There are 29 complete datasets, and 1 abbreviated run. The total size of the EEG files is 50.63 GB, representing 22.87 hours of EEG recording.

2.1.3.4 ARL_BCIT_SpeedControl Dataset

A total of 63 datasets were collected across 32 recording sessions, from 32 unique subjects recruited via the Teledyne laboratory. The subject ID can be used to identify data from subjects that participated in multiple experiments within program task 2.

There are 58 complete datasets, and 5 abbreviated runs. The total size of the EEG files is 99.32 GB, representing 44.76 hours of EEG recording.

2.1.3.5 ARL_BCIT_AuditoryCueing Dataset

A total of 34 datasets were collected across 17 recording sessions, from 17 unique subjects recruited via the Teledyne laboratory. The subject ID can be used to identify data from subjects that participated in multiple experiments within program task 2.

There are 34 complete datasets. None were abbreviated or interrupted. The total size of the EEG files is 57.85 GB, representing 26.00 hours of EEG recording.

2.1.3.6 ARL_BCIT_MindWandering Dataset

A total of 60 datasets were collected across 21 recording sessions, from 21 unique subjects recruited via the Teledyne laboratory. The subject ID can be used to identify data from subjects that participated in multiple experiments within program task 2.

There are 60 complete datasets. None were abbreviated or interrupted. The total size of the EEG files is 69.26 GB, representing 30.05 hours of EEG recording.

2.1.3.7 ARL_BCIT_RSVPBaseline Dataset

A total of 27 datasets were collected across 27 recording sessions, from 27 unique subjects recruited via the SAIC laboratory. The subject ID can be used to identify data from subjects that participated in multiple experiments within program task 3.

There are 23 complete datasets, and 4 that were interrupted by a system failure, but resumed. The total size of the EEG files is 271.20 GB, representing 31.83 hours of EEG recording.

2.1.3.8 ARL_BCIT_ RSVPExpertise Dataset

A total of 51 datasets were collected across 51 recording sessions, from 10 unique subjects recruited via the SAIC laboratory. Each subject was schedule to have data collected on 5 different days, although 1 subject participated in 6 RSVPExpertise sessions. The subject ID can be used to identify data from subjects that participated in multiple experiments within program task 3.

There are 43 complete datasets, and 8 that were interrupted by a system failure, but resumed. The total size of the EEG files is 493.10 GB, representing 59.72 hours of EEG recording.

2.1.3.9 ARL_BCIT_BasicGuardDuty Dataset

A total of 21 datasets were collected across 21 recording sessions, from 21 unique subjects recruited via the SAIC laboratory. The subject ID can be used to identify data from subjects that participated in multiple experiments within program task 3.

There are 20 complete datasets, and 1 that was interrupted by a system failure, but resumed. The total size of the EEG files is 164.80 GB, representing 19.07 hours of EEG recording.

2.1.3.10 ARL_BCIT_ AdvancedGuardDuty Dataset

A total of 27 datasets were collected across 27 recording sessions, from 27 unique subjects recruited via the SAIC laboratory. The subject ID can be used to identify data from subjects that participated in multiple experiments within program task 3.

There are 25 complete datasets, and 2 that was interrupted by a system failure, but resumed. The total size of the EEG files is 191.00 GB, representing 23.87 hours of EEG recording.

2.2 High Definition Cognition (HDCOG)

Technological advances can provide revolutionary capabilities, but may also go beyond Soldier cognitive capabilities, limiting how effectively advanced capabilities can be used. Design and integration of advanced systems and methods that account for how a Soldier's brain works would allow the matching of Soldier capabilities with system capabilities, including the use of real-time measures of cognition in systems design. However, traditional methods of cognitive performance assessment cannot always provide the objective, real-time understandings of Soldier cognition needed to obtain the target performance objectives. Neither the technologies to allow us to understand how a Soldier's brain works in operational environments, nor the techniques to integrate such understandings into system design have been fully developed. Yet emerging methods and technologies have been advancing rapidly in recent years and thus are becoming ready for assessment and study in concert with exemplar objective systems, to include advanced crew interfaces.

To meet these challenges, the High-Definition Cognition (HD-Cog) in Operational Environments Army Technology Objective (Research) (ATO-R) has been initiated as part of the U.S. Army Research Laboratory's Strategic Research Area in Neuroscience, with the goal of enabling more objective, direct, and higher-resolution assessment of Soldier cognitive processes to improve system design. Within this context, two of the significant focus areas of the HD-Cog ATO-R are the development and validation of operationally relevant metrics of Soldier cognitive function and techniques to use such metrics in improving systems design. Specifically, the efforts under this task order will 1) Integrate necessary hardware and software solutions for the acquisition of neurocognitive data sufficient to enable metrics development and experimentation; 2) Investigate techniques for integrating operationally-relevant cognitive metrics into neuroscience-based designs to enhance Soldier-system performance; 3) Support the design, development, and integration of an experimental scenario for use in joint ARL-TARDEC experimentation aimed at deriving and validating operationally-relevant neurocognitive metrics; 4) Transition and integrate software provided by ARL research partners for a prototype, multi-screen display for use in experimentation to develop an adaptive, attention-centric information interface based upon physiological and/or behavioral inputs; and 5) Support ongoing development and testing of an initial software implementation of a system that integrates information about mission, environment, and Soldier cognitive state to offload the task allocation process from the commander of a vehicle crew, balance workload among crew members, and focus the efforts of the crew on the most critical mission tasks.

The efforts of this task fall within the technical areas of the Cognition and Neuroergonomics (CaN) Collaborative Technology Alliance (CTA). The purpose of the CaN CTA is to: 1) extend the range and depth of known principles of and established methods for performing online estimation of the cognitive state, event appraisal, and behavioral intent of Army operators of complex systems in military-relevant, non-laboratory environments and 2) explore how the principles and methods thus established might best be used to design individualized real-time systems to improve situational awareness and decision-making under stress. The CaN CTA is divided into three major technical Thrust Areas (TA): 1) Neurocognitive Performance, 2) Advanced Computational Approaches, and 3) Neurotechnologies. First, the integration of hardware and software solutions supports all of the CaN CTA TAs inasmuch as these efforts underlie the abilities to develop metrics and methods for high-level experimentation. Second, investigation of cognitive metrics and the development of the experimental scenario for metrics development and validation supports both TA 1 and TA 2. Third, the goal of the adaptive information interface development is to take into account operator attentional state and underlying visual processing capabilities for improving event notification, also consistent with the focus of TA 1 and TA 3. Finally, the goal of the task allocation system development is to provide a system architecture that enables the application of operationally relevant metrics of Soldier cognition to improve vehicle crew performance by intelligently allocating tasks among the crewmembers, supporting the overall objectives of the CaN CTA.

2.2.1 HDCOG Program Summary

Data collection efforts within the HDCOG program consisted of a series of experiments conducted using 6-DOF Ride Motion Simulator (RMS) platform within the Ground Vehicle Simulation

Laboratory (GVSL) at the U.S. Army Tank and Automotive Research, Development, and Engineering Center (TARDEC) at the Detroit Arsenal in Warren, MI.

The naming convention for experiments conducted at GVSL is “TARDEC Experiment n” (or TXn”, where n is a sequential numeric identifier. HDCOG experiments include TX14, TX15, TX16, and TX17.

The simulation scenarios (target placement and planned vehicle intervention events) were the same for TX14 and TX15.

The objectives of the program were to 1) Integrate necessary hardware and software solutions for the acquisition of neurocognitive data; 2) Investigate techniques for integrating operationally-relevant cognitive metrics into systems design; 3) Support the design, development, and integration of a simulation scenario for use in experimentation to develop and validate operationally-relevant cognitive metrics; 4) Design and develop a prototype display for use in experimentation to develop an attention-centric information interface; and 5) Support ongoing development and testing of software for a proof-of-concept demonstration version of a task allocation system for vehicle crew workload management.

2.2.1.1 ARL HDCOG TARDEC Experiment 14 (ARL_HDCOG_TX14)

This Army’s transition to a leaner, more agile and rapidly deployable force requires the advent of autonomous technologies and systems, and more reliance on computers and machines. This move from traditional warfare to Future Combat Systems (FCS) represents a shift in the human role, as well. Technological advancement has made it so that the role of the user has been transformed from active controller to system monitor and manager, intervening only in the case of a problem. As such, the soldier’s dependency on robotics technologies, tele-operations, indirect driving and autonomy is expected to increase significantly. Additionally, although semi-autonomous driving technologies have proven beneficial in aggregate measures of local area awareness (i.e., target/threat detection) and vehicle control, it is important to understand the situational trade-offs between local area awareness and vehicle control, as situational trade-offs provide the basis for developing dynamic task allocation within crewstations.

2.2.1.1.1 Protocol

The research protocol information for this study is as follows:

> Organization: ARL/TARDEC
> Protocol Number: ARL-20098-09021
> Protocol Name: “The Physiological Basis of Local Area Security and Semi-Autonomous Driving”

2.2.1.1.2 Location

This study was conducted at TARDEC's Ride-Motion Simulator (RMS) in Warren, MI.

2.2.1.1.3 Subjects

A total of 21 subjects were recruited for this study. There is usable data from 20 subjects.

2.2.1.1.4 Apparatus

Subjects performed experiment tasks using a simulated crew-station mounted on the GVSL RMS. The vehicle simulation was simulated using high-fidelity vehicle modeling software called SimCreator (Real Time Technologies; Dearborn, MI), along with a custom-designed distributed system to integration the Crewstation interface, Intelligent Systems Behavior Simulator (ISBS), graphics processing for the simulation environment, EEG system, eye-tracking system, and the data logging components (see Figure 2-10).

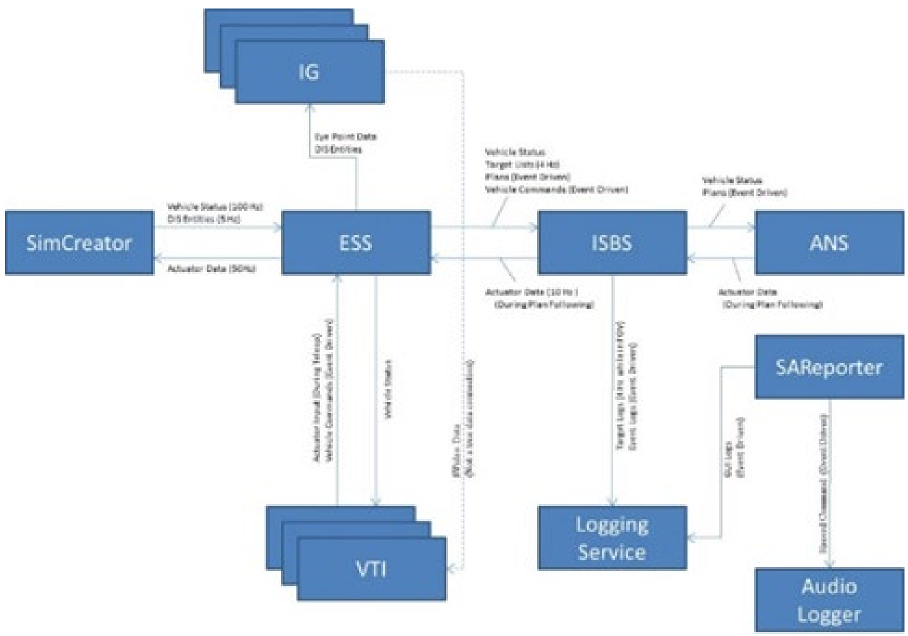


Figure 2-10. TX14 Data Collection System Architecture

The crew-station featured touch-screen control of an autonomous mobility system (AMS). Subjects could engage the system for autonomous movement, or disengage the AMS, and use a joystick for drive-by-wire vehicle control. The crewstation interface (see Figure 2-11) also contained a target reporting and classification mechanism, and video portals for situational awareness.

Figure 2-11. TX14 Crewstation Interface

The simulation environment utilized the "Desert Metro" terrain database, which was comprised of 6 stitched tiles to model a large city with roads, buildings, signs, and other features of a populated area. OpenFlight models for humans and vehicles were placed in the environment, following the Distributed Interactive Simulation (DIS) protocol, to create 4 unique "scenarios", each of which contained 20 events for the target detection task, and 3 events for the vehicle intervention task. They were intended to be statistically equivalent, however, initial analysis showed scenario 3 to be an outlier. Thus, the data collected against scenario 3 was removed from the final analyses.

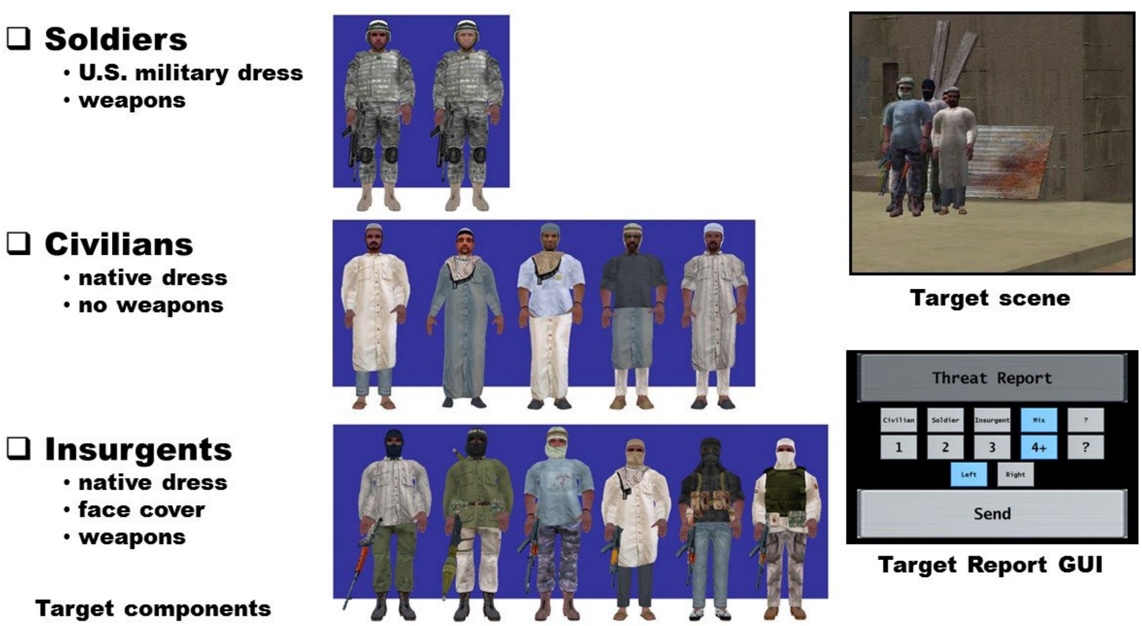
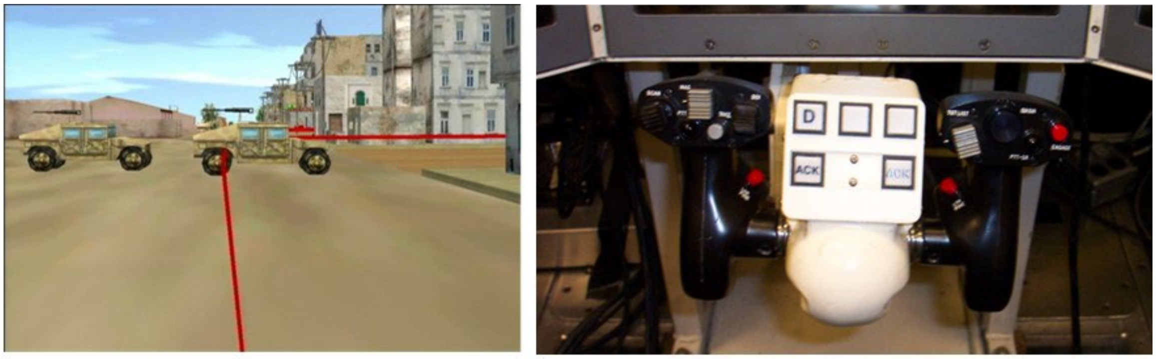


Figure 2-12. TX14 & TX15 Target Detection and Reporting



Road blockage through route **Crewstation Steering Device**

Figure 2-13. TX14 & TX15 Vehicle Intervention

The vehicle state data sampling rate was 100 Hz, logged along with crew-station interactions (i.e., subject behaviors) to aid post-mission analysis. Physiological measurements were also collected, with subjects wearing a BioSemi 64 (+8) EEG channel system with 4 eye and 2 mastoid channels recorded, with a sampling rate of 256 Hz. Eye Tracking data was collected using a faceLAB 4 system, with a 60 Hz sampling rate.

2.2.1.1.5 Demographics

No demographic data is provided with this dataset.

2.2.1.1.6 TX14 Tasks

There was one task specified for this experiment, which is described as follows:

> TX14:
> Duration: ~15 minutes
>
> Environment: Urban environment, representative of a middle-eastern city, included buildings, roads, traffic signs, and background features such as vehicles, clutter, etc. Humans, both static and in-motion, were placed as targets. Targets emerged in a naturalistic manner (i.e., did not "pop up" or "teleport in"), by appearing from around corners or through doorways, or by gaining line-of-sight to them as the vehicle rounded corners, etc.
>
> Scenarios: 4 unique sets of target locations and attributes. Event attributes and associated simulation models used for each target are defined in the data specification spreadsheet that accompanies the data.
>
> Task requirements: Vehicle moves on a fixed route using the AMS through the environment while subjects performed target detection and reporting tasks (shown in

Figure 2-12. TX14 & TX15 Target Detection and Reporting). At 3 pre-defined locations on the route, the road will be blocked (see Figure 2-13) When this occurs, the subject was instructed to disengage the AMS and manually control the vehicle around the obstacle, then re-engage the AMS.

Condition A: Vehicle motion only

Condition B: Vehicle motion + "noise" (vibration)

Sessions: Each subject was scheduled to perform the TX14 task 8 times on the same day. They performed the task under Condition A and Condition B against each of the 4 scenarios, with the sequence counter-balanced across subjects,

2.2.1.1.7 POC

Gabriella Larkin (gabriella.b.larkin.civ@army.mil), ARL, Primary Investigator, Analysis
James A. Davis, ARL, Associate Investigator, Analysis
Victor Paul (victor.j.paul2.civ@army.mil), GVSL, Associate Investigator
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer

2.2.1.2 ARL HDCOG TARDEC Experiment 15 (ARL_HDCOG_TX15)

This Army's transition to a leaner, more agile and rapidly deployable force requires the advent of autonomous technologies and systems, and more reliance on computers and machines. This move from traditional warfare to FCS represents a shift in the human role, as well. Technological advancement has made it so that the role of the user has been transformed from active controller to system monitor and manager, intervening only in the case of a problem. As such, the soldier's dependency on robotics technologies, tele-operations, indirect driving and autonomy is expected to increase significantly. Additionally, although semi-autonomous driving technologies have proven beneficial in aggregate measures of local area awareness (i.e., target/threat detection) and vehicle control, it is important to understand the situational trade-offs between local area awareness and vehicle control, as situational trade-offs provide the basis for developing dynamic task allocation within crewstations.

2.2.1.2.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL/TARDEC
Protocol Number: ARL-20098-09021

Protocol Name: “The Physiological Basis of Local Area Security and Semi-Autonomous Driving”

2.2.1.2.2 Location

This study was conducted at TARDEC’s Ride-Motion Simulator (RMS) in Warren, MI.

2.2.1.2.3 Subjects

14 soldier subjects were recruited for this study.

2.2.1.2.4 Apparatus

Subjects performed experiment tasks using a simulated crew-station mounted on the GVSL RMS. The vehicle simulation was simulated using high-fidelity vehicle modeling software called SimCreator (Real Time Technologies; Dearborn, MI).

The crew-station featured touch-screen control of an autonomous mobility system (AMS). Subjects could engage the system for autonomous movement, or disengage the AMS, and use a joystick for drive-by-wire vehicle control. The crewstation interface also contained a target reporting and classification mechanism, and video portals for situational awareness.

The vehicle state data sampling rate was 100 Hz, logged along with crew-station interactions (i.e., subject behaviors) to aid post-mission analysis. Physiological measurements were also collected, with subjects wearing a BioSemi 64 (+8) EEG channel system with 4 eye and 2 mastoid channels recorded, with a sampling rate of 256 Hz. Eye Tracking data was collected using a faceLAB 4 system, with a 60 Hz sampling rate.

2.2.1.2.5 Demographics

No demographic data is provided with this dataset.

2.2.1.2.6 TX15 Tasks

There were two tasks specified for this experiment, which are described as follows:

TX15:
Duration: ~15 minutes

Environment: Urban environment, representative of a middle-eastern city, included buildings, roads, traffic signs, and background features such as vehicles, clutter, etc. Humans, both static and in-motion, were placed as targets. Targets emerged in a naturalistic manner (i.e., did not "pop up" or "teleport in"), by appearing from around corners or through doorways, or by gaining line-of-sight to them as the vehicle rounded corners, etc.

Scenarios: 4 unique sets of target locations and attributes. They were intended to be statistically equivalent, however, initial analysis showed scenario 3 to be an outlier. Thus, the data collected against this scenario was removed from the formal analyses.

Task requirements: Vehicle moves on a fixed route using the AMS through the environment while subjects performed target detection and reporting tasks (shown in Figure 2-12. TX14 & TX15 Target Detection and Reporting). At 3 pre-defined locations on the route, the road will be blocked (see Figure 2-13) When this occurs, the subject was instructed to disengage the AMS and manually control the vehicle around the obstacle, then re-engage the AMS.

Condition A: Crewstation video and vehicle motion synchronized

Condition B: Crewstation video and vehicle motion de-coupled; visual feedback briefly lags behind motion

Sessions: The first 5 subjects were scheduled to perform the TX15 task 8 times on the same day. They performed the task under Condition A and Condition B against scenarios 1, 2, 3, and 4, with the sequence counter-balanced across subjects. The last 9 subjects were scheduled to perform the TX15 task 6 times on the same day. They performed the task under Condition A and Condition B against scenarios 1, 2, and 4, with the sequence counter-balanced across subjects.

<u>TX15 Oddball</u>:
Duration: ~15 minutes

Environment: The motion platform was turned off for this task, although the vehicle did move through the simulated environment during the test. A view of the (changing) urban environment was displayed in the background, and Gabor patch images were displayed as overlays on the center screen.

Scenarios: Standard stimulus images were displayed 88% of the time and oddball images were displayed 12% of the time, as shown in Figure 2-14. The standard stimulus consisted of a Gabor grating with a green square in the center, and the rare stimulus was a Gabor grating with a blue circle in the center. Each stimulus was presented for 250ms in the center of the middle crewstation display and separated by 1-1.5s.

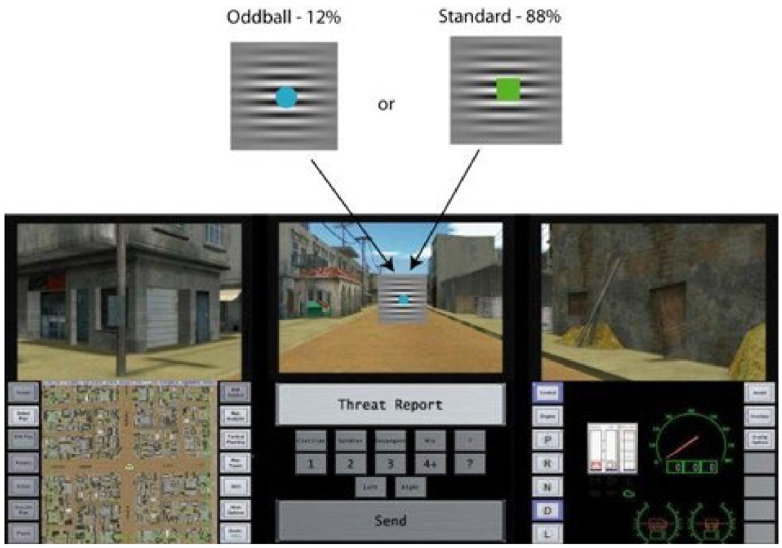


Figure 2-14. TX15 Visual Oddball Task

Task requirements: Subjects experienced a visual oddball paradigm, which required them to discern between oddball images (targets) and other images, and press different buttons in each case.

Sessions: The last 9 subjects that were scheduled to perform the TX15 task, also performed the TX15_Oddball task, as the last task of the session.

2.2.1.2.7 POC

Gabriella Larkin (gabriella.b.larkin.civ@army.mil), ARL, Primary Investigator, Analysis
James A. Davis, ARL, Associate Investigator, Analysis
Anthony J. Ries (anthony.j.ries2.civ@army.mil), ARL, Associate Investigator, Analysis
Kelvin S. Oie, ARL, Associate Investigator, Analysis
Victor Paul (victor.j.paul2.civ@army.mil), GVSL, Associate Investigator
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer

2.2.1.3 ARL HDCOG TARDEC Experiment 16 (ARL_HDCOG_TX16)

Commanders of military vehicles are responsible for allocating the tasks of a mission plan to the crewmembers who are operating the vehicle. Within the US Army, the classic approach has been to define a role for each crewmember and to predefine the types of tasks that should be assigned to each role. This approach has made it possible to design a different task-specific crewstation and to train crewmembers for each role. During a mission, each task maps to exactly one crewmember, so there is no confusion about who should perform each task.

The US Army is currently developing a System of Systems (SoS) containing manned vehicles, unmanned vehicles, ground sensors, and Soldiers all working together through an integrated network. One of the objectives is to increase Soldier survivability in ground vehicle operations. This objective is expected to be accomplished through increased mobility and self-contained operations as well as increased reliance on complex information networks, while, at the same time, minimizing crew size. The combination of an increased number of tasks for each member and fewer crewmembers means that it is essential to manage the task allocation process intelligently.

Several new technologies are currently being developed to help maximize Soldier performance in vehicles. TARDEC is developing a Warfighter Machine Interface (WMI) that allows each crewmember to perform almost any mission task. The U.S. Army Human Research and Engineering Directorate (HRED) is developing a suite of sensors to measure the physiological and cognitive states of Soldiers in operational environments.

This experiment incorporates two of the underlying computational components that are needed to fuel the technical development of the HRED sensor suite. The first includes modifications in the simulation design algorithms to increase the realism of the task and provide more interaction with and control of the simulated task environment to the participants. The second investigates the feasibility of existing algorithms to successfully classify the participant's mental state, and in particular, to discriminate times of high and low fatigue and times of high task difficulty compared to times of low task difficulty. Together, these components will enable the development of metrics to assess physiological and cognitive states of Soldiers to maximize Soldier performance in operational environments.

2.2.1.3.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL/TARDEC
Protocol Number: ARL-20098-10051 (ARL 10-051)
Protocol Name: "Dynamic Classification of Soldier State"

2.2.1.3.2 Location

This study was conducted at TARDEC's Ride-Motion Simulator (RMS) in Warren, MI.

2.2.1.3.3 Subjects

A total of 14 subjects were recruited for this study to perform the role of Vehicle Commander (VC), which was the subject under study (and wearing the EEG system). An additional 14 participants served as the Driver for the VC, controlling the simulated vehicle and communicating with the VC via radio headset. The driver was not instrumented or measured during the experiment.

2.2.1.3.4 Apparatus

Subjects performed experiment tasks using a simulated crew-station cab mounted on the GVSL RMS. The vehicle simulation was simulated using high-fidelity vehicle modeling software called SimCreator (Real Time Technologies; Dearborn, MI), along with a custom-designed distributed simulation system to integrate the Crewstation interface, Scenario Populator, Event Server (recognized trip lines), graphics processing for the simulation environment, EEG system, eye-tracking system, and the data logging components. While the system architecture diagram shown in Figure 1-5 indicates the planned used of a respiration belt and GSR measurement, those devices did not make it into the final system configuration.

The Vehicle Commander (VC) crew-station interface, shown in Figure 2-15, featured touch-screen control of 360-degree camera suite bringing continuous video from up to 4 of 6 total cameras at one time through the use of multiple video portals. There was also an overview map indicating, in real-time, vehicle position and heading within the environment, as well as roadways and buildings.

Figure 2-15. TX16 Crewstation Interface

The VC also wore a radio headset with microphone, and used a plunger button to simulate keying the mic before speaking. There were two separate radio networks over which simulated communications could be broadcast, allowing for overlap of messages.

The vehicle state data sampling rate was 100 Hz, logged along with crew-station interactions (i.e., subject behaviors) to aid post-mission analysis. Physiological measurements were also collected, with subjects wearing a BioSemi 64 (+8) EEG channel system with 4 eye and 2 mastoid channels recorded, with a sampling rate of 256 Hz. Eye Tracking data was collected using a faceLAB 4 system, with a 60 Hz sampling rate.

2.2.1.3.5 Demographics

Demographic information provided with this dataset includes subject ID, gender, age, dominant hand, vision (normal or corrected to normal), and hearing (normal or corrected to normal).

2.2.1.3.6 TX16 Tasks

There were two tasks specified for this experiment, which are described as follows:

<u>TX16</u>:
Duration: ~20 minutes

Environment: Urban environment, representative of a middle-eastern city, included buildings, roads, traffic signs, and background features such as vehicles, clutter, etc. Humans, both static and in-motion, were placed in the environment. Targets to be reported included Iraqi Army soldiers (IA's), loose weapons in the environment, and individuals fitting the description of "BOLO" alert communications.

Scenarios: 6 unique combinations of checkpoints, audio stimuli, and visual stimuli were generated to serve as distinct but statistically equivalent scenarios.

Task requirements: Fulfill the role of Vehicle Commander, and assume all relevant responsibilities, including:

1. Route planning (prior to the mission)
2. Arriving at 3 checkpoints at specified times
3. Provide navigation commands to driver
4. 360-degree Local Situational Awareness (LSA)
    a. Report uniformed local armed forces, illicit weapons, and BOLO's to FOB BC
    b. Sample Alert: "*BOLO for a female civilian wearing brown clothing and a scarf with a blue band*"
    c. Sample Report from subject: "*Mother Hen this is Blue4: the civilian with a blue band on her scarf has been spotted at 4 o'clock near the city entrance*"
5. Monitoring and responding to audio communications over two radio networks
    a. Subject responded when addressed by their call sign, "BLUE 4", and when messages were broadcast to "ALL CON"
6. Report low wires and/or overpasses for future convoy
    a. Subjects reported "Approaching low wires: 3, 2, 1, MARK"

Note: A detailed description of subject tasks is provided in the "TX16 Overview" PowerPoint file in the "additional information" folder with the dataset. This file was used for subject training. This file should be referenced for a complete understanding of what subjects experienced, and what information is available.

Sessions: Subjects were scheduled to perform the TX16 task 6 times on the same day; 1 run each using 6 different scenarios, with the sequence counter-balanced across subjects.

TX16 Auditory Versus Visual Discrimination:
Duration: ~15 minutes

Environment: The motion platform was turned off for this task, although the vehicle did move through the simulated environment during the test. A view of the (changing) urban environment was displayed in the background, along with human targets.

Scenarios: Presentation of stimuli was segregated for this task. During the portion of the drive from the FOB to the city edge, audio messages played, and no visual stimuli were presented. After reaching the city, the audio messages ceased, and visual stimuli were presented in the form of targets (Iraqi Soldiers) and non-targets (civilians and non-human objects).

Task requirements: Subjects were required to respond to audio messages which were addressed to them (via call sign), and to report targets viewed in the city.

Sessions: Subjects were scheduled to perform the TX16 AuditoryVsVisual task as the last task in a session that included all runs of the basic TX16 task.

2.2.1.3.7 POC

Brent J. Lance (brent.j.lance.civ@army.mil), ARL, Primary Investigator, Analysis
Jean M. Vettel (jean.m.vettel.civ@army.mil), ARL, Primary Investigator, Analysis
Victor Paul (victor.j.paul2.civ@army.mil), GVSL, Associate Investigator
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer
Mike Dunkel (mdunkel@omi.com), DCS Corp., Data Engineer

2.2.1.4 ARL HDCOG TARDEC Experiment 17A (ARL_HDCOG_TX17A)

The purpose of this project is to investigate the reliability and generalizability of a neural fatigue-based driver performance prediction methodology and a neural workload-based driver performance prediction methodology on Army-relevant simulated driving tasks. The protocol aims to compare the ability of existing algorithms to dynamically classify a participant's fatigue state during the simulated mission. The protocol also aims to compare the ability of existing algorithms to dynamically classify a participant's mental state at varying levels of task difficulty for the participants throughout the simulated mission. For each experimental session, one Soldier participant will be recruited to perform two driving scenarios, one low-activity for evaluation with the fatigue-based monitor, and one high-activity for evaluation with the workload-based monitor. The majority of the participants will be active duty Soldiers recruited by TARDEC and Joint Program Office (JPO) Mine Resistant Ambush Protected (MRAP). The Soldiers will fly to Warren, MI to be tested in the Ground Vehicle Simulation Laboratory (GVSL) located at TARDEC at the Detroit Arsenal.

2.2.1.4.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL/TARDEC
Protocol Number: ARL-20098-10051
Protocol Name: “Army Relevant EEG Based Driver Performance Prediction”
Contract: W911NF-10-D-0002-0003

2.2.1.4.2 Location

This study was conducted at TARDEC’s Ride-Motion Simulator (RMS) in Warren, MI.

2.2.1.4.3 Subjects

A total of 11 subjects were recruited for this study, and 11 sessions of 2 runs per subject yielded 20 recordings of usable data.

2.2.1.4.4 Apparatus

Subjects performed the driving task using a simulated cab and crew-station mounted on the GVSL RMS, with a yoke for steering, and pedals for acceleration and braking. The vehicle simulation was simulated using high-fidelity vehicle modeling software called SimCreator (Real Time Technologies; Dearborn, MI). The simulated environment featured a racetrack design road, depicted in Figure 2-16.

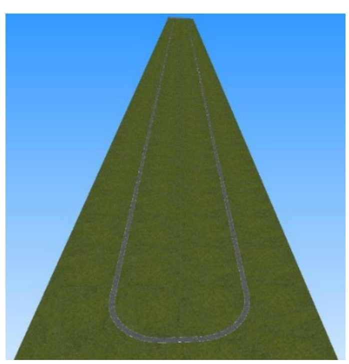


Figure 2-16. TX17A Simulated Roadway

The vehicle state data sampling rate was 100 Hz, logged along with crew-station interactions (i.e., subject behaviors) to aid post-mission analysis. Physiological measurements were also collected, with subjects wearing a BioSemi 64 (+8) EEG channel system with 4 eye and 2 mastoid channels recorded, with a sampling rate of 256 Hz. Eye Tracking data was collected using a SmartEYE system, with a 60 Hz sampling rate.

2.2.1.4.5 Demographics

Demographic information provided with this dataset includes subject ID, gender, age, dominant hand, vision (normal or corrected to normal), and hearing (normal or corrected to normal).

2.2.1.4.6 TX17A Tasks

There was one task specified for this experiment, which is described as follows:

> TX17A:
> Duration: ~45 minutes
>
> Environment: Racetrack (loop) road within a visually sparse environment.
>
> Scenarios: Perturbation events were scheduled under the following conditions throughout each run:
>
> > Vehicle must not be currently experiencing a perturbation, AND must be within the lane boundaries for 8 seconds; then, a perturbation event is scheduled to begin at a randomly selected time between 0 and 2 seconds.
>
> Task requirements: Keep the vehicle within lane boundaries, correcting for perturbations as necessary, and otherwise remaining vigilant.
>
> Sessions: Subjects were scheduled to perform the TX17A task 2 times on the same day

2.2.1.4.7 POC

Brent J. Lance (brent.j.lance.civ@army.mil), ARL, Primary Investigator, Analysis
Victor Paul (victor.j.paul2.civ@army.mil), GVSL, Associate Investigator
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer

2.2.2 HDCOG Publications

Jason Metcalfe, Gabriella Larkin, Tony Johnson, Kelvin Oie, Victor Paul, Jams A. Davis. (2010) **Experimentation and evaluation of threat detection and local area awareness using advanced computational technologies in a simulated military environment**. *Proceedings of SPIE Vol. 7692* (SPIE, Bellingham, WA, 2010) 769209.
DOI: 10.1117/12.850516

Anthony Ries, Victor Paul, Marcel Cannon, Kelvin Oie (2010) **NEURAL RESPONSES TO COMMON AND RARE VISUAL EVENTS DURING A SIMULATED DRIVING SCENARIO**. Army Science Conference, 2010

Jean Vettel, Brent Lance, Chris Manteuffel, Matthew Jaswa, Marcel Cannon, Tony Johnson, Victor Paul, Kelvin Oie (2012) **Mission-based Scenario Research: Experimental Design and Analysis**, *Ground Vehicle Systems Engineering Symposium August 9-11*
http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwiw4eD1kMDUAhVlDZoKHY3EAwYQFggpMAA&url=http%3A%2F%2Fwww.dtic.mil%2Fget-tr-doc%2Fpdf%3FAD%3DADA556740&usg=AFQjCNHgYp090gtp1y48GbIHzxuC6tE6QQ&sig2=SEIUiSLVQpSsgnLSiy2GeQ

L. M. Merino, J. Meng, S. Gordon, B. J. Lance, T. Johnson, V. Paul, K. Robbins, J.M. Vettel, & Y. Huang (2013). **A bag-of-words model for task-load prediction from EEG in complex environments**. In *2013 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013 - Proceedings* (pp. 1227-1231). [6637846]
DOI: 10.1109/ICASSP.2013.6637846
http://ieeexplore.ieee.org/document/6637846/

2.2.3 HDCOG Datasets

A total of 363 datasets were collected across 82 recording sessions, using 82 unique subjects, within the 4 specified experiments.

The total size of the EEG files is 54.12 GB, representing 94.19 hours of EEG recording.

2.2.3.1 ARL_HDCOG_TX14 Dataset

A total of 147 datasets were collected across 20 recording sessions, from 20 unique subjects recruited via the GVSL.

The total size of the EEG files is 15.52 GB, representing 28.42 hours of EEG recording.

2.2.3.2 ARL_HDCOG_TX15 Dataset

A total of 91 datasets were collected across 14 recording sessions, from 14 unique subjects recruited via the GVSL. This total includes 88 complete datasets, and 3 abbreviated datasets.

The total size of the EEG files is 10.38 GB, representing 18.68 hours of EEG recording.

2.2.3.3 ARL_HDCOG_TX15_Oddball Dataset

A total of 8 datasets were collected across 8 recording sessions, from 8 unique subjects recruited via the GVSL as part of the TX15 experiment. The “oddball” experiment was included as the final task performed by TX15 subjects in their sessions.

The total size of the EEG files is 1.48 GB, representing 0.82 hours of EEG recording.

2.2.3.4 ARL_HDCOG_TX16 Dataset

A total of 81 datasets were collected across 14 recording sessions, from 14 unique subjects recruited via the GVSL. This total includes 68 complete datasets, and 13 abbreviated datasets.

The total size of the EEG files is 26.16 GB, representing 14.52 hours of EEG recording.

2.2.3.5 ARL_HDCOG_TX16_AuditoryVsVisual Dataset

A total of 13 datasets were collected across 13 recording sessions, from 13 unique subjects recruited via the GVSL as part of the TX16 experiment. This special comparison of segregated stimuli types (audio only first, then visual only) was included as the final task performed by TX16 subjects in their sessions.

The total size of the EEG files is 3.49 GB, representing 1.94 hours of EEG recording.

2.2.3.6 ARL_HDCOG_TX17A Dataset

A total of 20 datasets were collected across 11 recording sessions, from 11 unique subjects recruited via the GVSL.

The total size of the EEG files is 10.94 GB, representing 15.95 hours of EEG recording.

2.3 Institute for Collaborative Biotechnologies (ICB)

The Institute for Collaborative Biotechnologies (ICB) is a University Affiliated Research Center (UARC) primarily funded by the United States Army. Headquartered at the University of California, Santa Barbara (UCSB) and in collaboration with MIT, Caltech and industry partners, ICB's interdisciplinary approach to research aims to enhance military technology by transforming biological systems into technological applications.

2.3.1 ICB Program Summary

The ICB's research aim is to model biological mechanisms for use in military materials and tools. Quoting Army Research Office program manager Robert Campbell, "The inspiration for the ICB comes from the fact that biology uses different mechanisms to produce materials and integrated circuits for high-performance sensing, computing and information processing, and actuation than are presently used in human manufacturing." Much research is focused on evaluating biomolecular sensors, bio-inspired materials and energy, biodiscovery tools, bio-inspired network science, and cognitive neuroscience through the disciplines of cellular and molecular biology, materials science, chemical engineering, mechanical engineering, and psychology.

2.3.1.1 RSVP Cognitive Technology Threat Warning System (CT2WS) Experiment (ARL_ICB_CT2WS)

The goal of this research effort is to investigate and define perceptual and cognitive processes and neural networks used to allocate visual attentional resources while multi-tasking. This information can then be exploited to enhance adaptive interface technologies that will facilitate superior allocation of cognitive resources and minimize distraction.

Indirect vision and multiple screen displays are an integral part of the Army's future systems. These advances in technology represent both the opportunity to further military (and industrial) capabilities and an all the more critical need to ensure the cognitive compatibility of these new technological feats. For example, multiple screen displays possess obvious advantages, such as the potential for displaying more information. The caveat to this opportunity is that while more information may result in enhanced situation awareness (SA), increased information presentation requires a higher degree of monitoring and vigilance, and is associated with a higher cognitive load. Vigilance as a cognitive process is of particular importance in the military setting, from the radar screens of fighter pilots in the sky to infantrymen's local area awareness on the ground. These displays provide a unique opportunity in cognitively compatible design, to alleviate the burden that vigilance places on neuro-cognitive processes. Therefore, a focus on increased SA capabilities must be tempered by an understanding of human neuro-cognitive limitations. Using human attention as a central construct and organizing principle in the design and enhancement of computation, communication, and other information systems can result in superior information management and an overall enhancement of user capabilities. One potential application for such a system would involve the classification of attentional/cognitive state based on non-invasive physiological measurements. The system can then dynamically assign and reassign different tasks based upon the physiological assessment of user state.

The general goal of this research is to evaluate the effect that both central and peripheral cues used to alert the user to secondary events have on serial multi-tasking. Centrally presented cues may provide a way in which to support and augment human attentional processing. As attention is crucial in the allocation of cognitive resources, the objective of this research is to assess the effect of centrally presented cues on the allocation of cognitive resources during serial multi-tasking.

An additional consideration in the development of adaptive displays pertains to the interfacing of neural signals with the computer system. There are many potential applications for such interfaces, and numerous aspects of neural activity that can be harnessed. This study will utilize EEG recordings. Prior research has demonstrated that using EEG measures in a rapid serial visual presentation paradigm, the occurrence of a threat stimulus could modulate perceptual event related potentials (ERP) components. Therefore, a secondary objective of the proposed research is to replicate these findings and to determine whether there is a difference in the isolation and reliable identification of brain signals associated with target detection between the two cuing conditions, and whether the two conditions affect the accuracy and reliability of target signature detection derived from EEG recordings.

2.3.1.1.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL
Protocol Number:
Protocol Name: "Impact of Advanced Display Features on Operator Performance"

2.3.1.1.2 Location

This study was conducted at ARL HRED, Aberdeen Proving Ground, MD.

2.3.1.1.3 Subjects

A total of 17 subjects were recruited for this study, from among local researchers, engineers, military personnel, and other staff at ARL HRED.

2.3.1.1.4 Apparatus

This project required the use of 3 Shuttle PCs, deployed as a distributed system. Each computer hosted custom software to manage the 3 main tasks within the simulation system:

1. RSVP task (primary task, center display)
2. Target Detection task (secondary task, left display)
3. Formation Deviation task (secondary task, right display)

The computers were connected to each other via network through a hub. The computer hosting the primary task also had a parallel port connection to the EEG system, for sending event codes used as marker data, written to the EEG data stream.

Subjects were instrumented with a BioSemi 64 (+8) EEG channel system with 4 eye and 2 mastoid channels recorded, with a sampling rate of 512 Hz.

2.3.1.1.5 Demographics

No demographic data is provided with this dataset.

2.3.1.1.6 RSVP CT2WS Tasks

There was one task (set of subject responsibilities) specified for this experiment, comprised of 3 separate activities. Each component is described as follows:

1. <u>RSVP activity (primary task, center display)</u>
   Subjects were instructed to make a manual button press with their dominant hand when they detected a target (person or vehicle) within a series of CT2WS video clips presented in an RSVP paradigm. Video clips consisted of five consecutive images, each 100 ms in duration; each video clip was presented for 500 ms. There was no interval between videos such that the first frame was presented immediately after the last frame of the prior video. If a target appeared in the video clip, it was present on each 100-ms image. The distractor-to-target ratio was 90/10. RSVP sequences were presented in 2-min blocks after which time observers were given a short break. A 'blink' message was displayed on the middle screen every 10 seconds. Sample images are shown in Figure 2-17.

2. <u>Target Detection activity (secondary task, left display)</u>
   This task uses images from the previous RSVP block, so it must only occur after the first RSVP block is complete. Upon receiving a cue to perform this task, the subject presses the "Start" button (center of screen), then receives a series of individual images from the set previously presented on the center screen. The number of images selected for evaluation is 20. The Distractor/Target ratio is 50/50. The subject performs a more thorough analysis of each image and reports "Target" or "No Target" for each one and presses "Submit". Images are advanced to the next one upon each response. There's no time limit to evaluate an image & respond.

3. <u>Formation Deviation activity (secondary task, right display)</u>
   Upon receiving a cue to perform this task, the subject presses the "Start" button (center of screen), and then the formation misbehavior is displayed. The subject reports on

periodic events related to coordinated asset movement (i.e. a formation monitoring task), determining the nature of formation violations based on the color and shape of the indicator, and indicating a corrective action. The subject selects the button that represents the correction command for the vehicle:

Red Triangle indicator = press “Veer Right” button
Green Triangle indicator = press “Veer Left” button
Green Circle indicator = press “Slow Down” button
Red Circle indicator = press “Speed Up” button

Triangles point in the direction of the misbehavior.
The subject presses the “Submit” button to complete the task.

The sequence of the task variants is as follows:

RSVP CT2WS Baseline condition:
Duration: ~5 minutes

Task requirements: Subject performed the RSVP activity only.

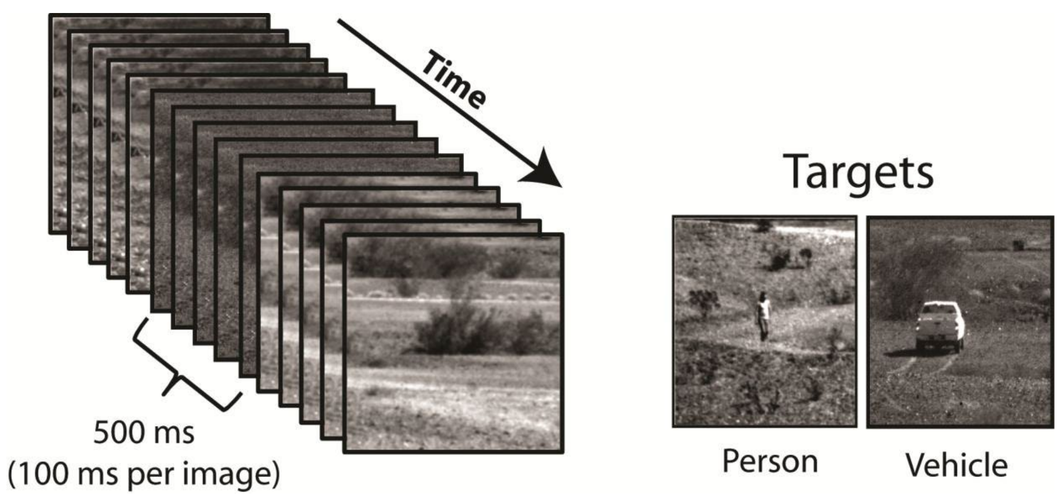


Figure 2-17. RSVP CT2WS Target Images & Display Timing

RSVP CT2WS Multitasking Condition 2 Central Cueing:
Duration: ~9 minutes

Task requirements: Subject performed the primary task, and attended to secondary tasks as alerted through a color overlay on the center screen, as shown in Figure 2-18. A red overlay indicated a need to work on the Target Detection task, and the green overlay indicated a need to work on the Formation Deviation task.

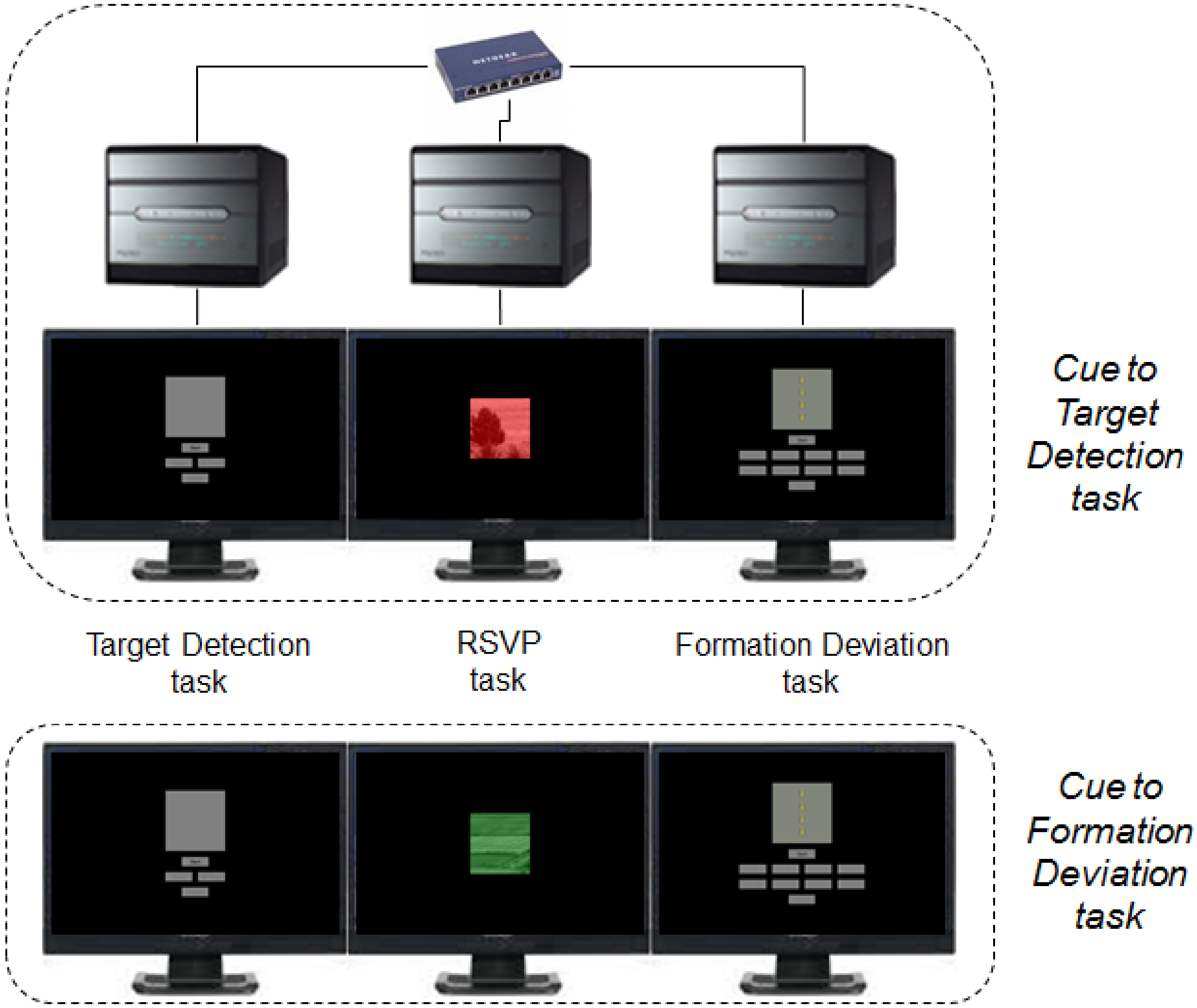


Figure 2-18. RSVP CT2WS Central Cueing

RSVP CT2WS Multitasking Condition 2 Peripheral Cueing:
Duration: ~9 minutes

Task requirements: Subject performed the primary task, and attend to secondary tasks as alerted through a color overlay on the side display corresponding with the activity that needed attention. A red overlay on the left display indicates a need to work on the Target Detection task, and the green overlay on the right display indicated a need to work on the Formation Deviation task. Examples of each task cue are shown in Figure 2-19.

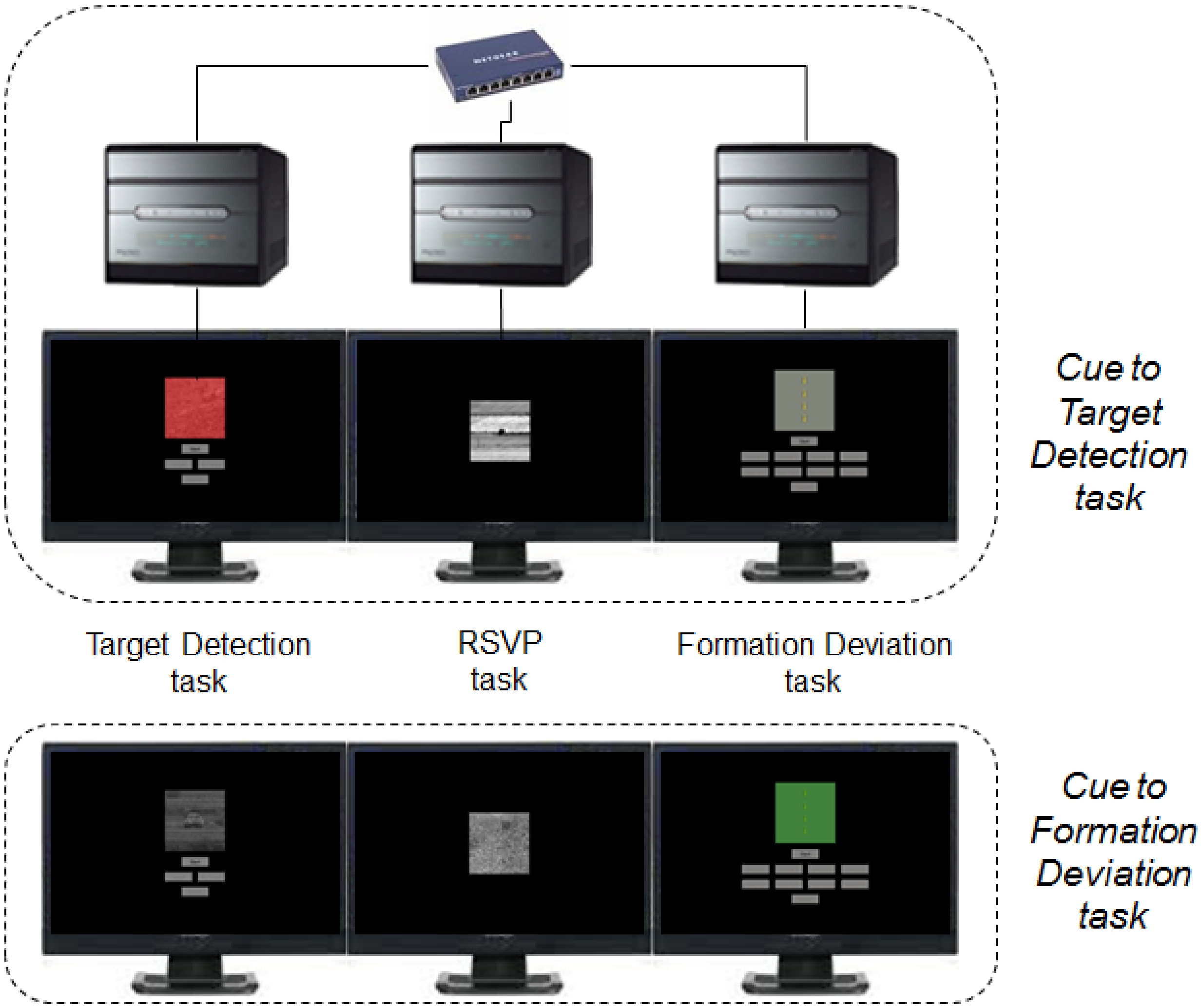


Figure 2-19. RSVP CT2WS Peripheral Cueing

Sessions: Each subject was scheduled to perform the 3 components of the RSVP CT2WS task on the same day, in a single recording session. The Baseline was performed first, followed by the 2 multi-tasking conditions with the condition sequence counter-balanced across subjects.

2.3.1.1.7 POC

Gabriella Larkin (gabriella.b.larkin.civ@army.mil), ARL, Primary Investigator, Analysis
Anthony J. Ries (anthony.j.ries2.civ@army.mil), ARL, Associate Investigator, Analysis
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer

2.3.1.2 ARL ICB RSVP Insurgent-Civilian (ARL_ICB_RSVP)

The RSVP Insurgent-Civilian experiment supports several analysis efforts in the areas of advanced algorithms and Brain-Computer Interaction (BCI).

2.3.1.2.1 Protocol

The research protocol information for this study is as follows:

> Organization: ARL/TARDEC
> Protocol Number: ARL-20098-09021
> Protocol Name: "The Physiological Basis of Local Area Security and Semi-Autonomous Driving"

2.3.1.2.2 Location

This study was conducted at ARL HRED, Aberdeen Proving Ground, MD.

2.3.1.2.3 Subjects

A total of 18 subjects were recruited for this study, although there is only usable data for 16 subjects.

2.3.1.2.4 Apparatus

A Dell Precision T7400 PC was used to host custom RSVP presentation and data collection software using E-Prime. Stimuli consisted of images that were 960x600 pixels, 96 dpi, and subtending 36.3 x 22.5.

Subjects were instrumented with a BioSemi 64 (+8) EEG channel system with 4 eye (EOG) channels and 2 mastoid channels recorded. The scalp electrodes were arranged in a 10-10 montage. The sampling rate was 1024 Hz.

2.3.1.2.5 Demographics

Demographic information provided with this dataset includes subject ID, gender, age, dominant hand, vision (normal or corrected to normal), caffeine intake, and alcohol intake.

2.3.1.2.6 RSVP Insurgent-Civilian Tasks

There was one task specified for this experiment, performed by each subject under 5 different conditions. The task is described as follows:

RSVP Insurgent-Civilian:
Duration: ~15 minutes

Environment: Urban environment, representative of a middle-eastern city, included buildings, roads, traffic signs, and background features such as vehicles, foliage, etc.

Stimuli: Images were presented for 500 ms (2 Hz) with no inter-stimulus interval. Images contained either a scene without any people (non-target) or a scene with a person holding a gun (target). Figure 2-20 and Figure 2-21 represent imagery used for this task containing sample RSVP stimuli and a detailed view of targets, respectively.

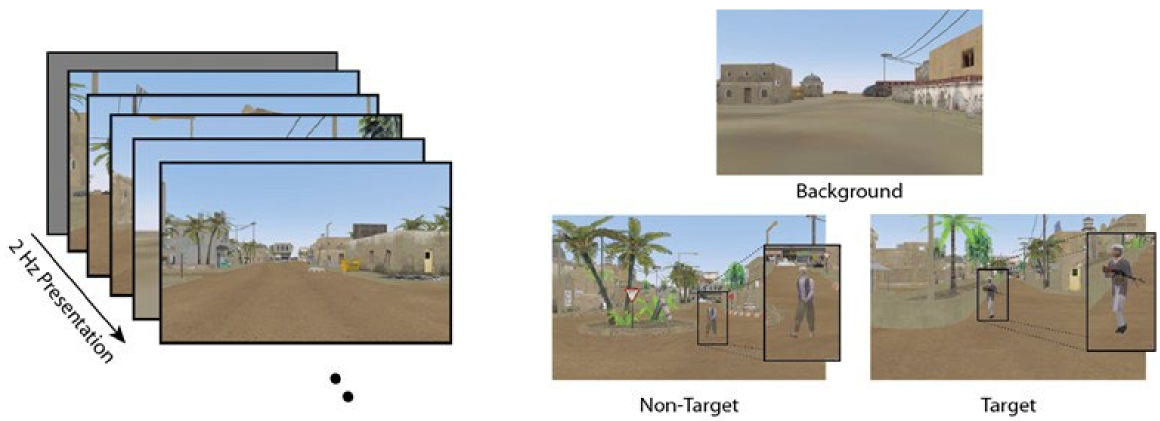


Figure 2-20. RSVP Insurgent-Civilian Images

A total number of 110 target images and 1346 non-target images were presented to each participant. Scenes in which a target appeared were also presented without the person in the non-target condition. All stimuli appeared within 6.5º of center of the monitor.

Figure 2-21. RSVP Insurgent-Civilian Targets

Task requirements: The goal of the task was to classify target images (humans with guns) from non-target images (humans without guns).

Condition B: Subjects performed a baseline version of the task involving 4 steps, 1-minute each:

Step 1: Subjects listened to a series of tones presented every second for one minute with no action

Step 2: Subjects blinked their eyes to the audio tones for one minute

Step 3: Subjects fixated on a cross on the computer screen in silence for one minute

Step 4: Subjects sat with their eyes closed for the one minute

Condition 1: Subjects performed the task while receiving stimuli containing *only* targets, and with a requirement to *only* count the number of targets.

Condition 2: Subjects performed the task while receiving stimuli containing *only* targets, and with a requirement to count the number of targets *and* respond to each with a button press.

Condition 3: Subjects performed the task while receiving stimuli containing targets *and* non-targets, and with a requirement to *only* count the number of targets

Condition 4: Subjects performed the task while receiving stimuli containing targets *and* non-targets, and with a requirement to count the number of targets *and* respond to each with a button press

Sessions: Subjects participated in a single session, and performed the RSVP task under all 5 conditions, with the condition sequence counter-balanced across subjects.

2.3.1.2.7 POC

Anthony J. Ries (anthony.j.ries2.civ@army.mil), ARL, Associate Investigator, Analysis
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer

2.3.2 ICB Publications

Anthony Ries, Gabriella Larkin (2013) **Stimulus and Response-Locked P3 Activity in a Dynamic Rapid Serial Visual Presentation (RSVP) Task.** ARL-TR-6314. January, 2013. http://www.dtic.mil/get-tr-doc/pdf?AD=ADA579452

Amar R. Marathe, Anthony J. Ries, Vernon J. Lawhern, Brent J. Lance, Jonathan Touryan, Kaleb McDowell, and Hubert Cecotti, (2015) **Effect of target and non-target similarity on neural classification performance: A boost from confidence**. *Frontiers in Neuroscience* 9:270.
DOI: 10.3389/fnins.2015.00270

http://journal.frontiersin.org/article/10.3389/fnins.2015.00270/full

Hubert Cecotti, and Anthony J. Ries, (2015). **Implication of non-stationarity in single-trial detection performance of event-related potentials**. IEEE-EMBC
http://uir.ulster.ac.uk/32324/1/paper02_ukci2015.pdf

Hubert Cecotti, Amar R. Marathe, and Anthony J. Ries, (2015). **Optimization of single-trial detection of event-related potentials through artificial trials**, IEEE TBME-01566-2014
DOI: 10.1109/TBME.2015.2417054
http://ieeexplore.ieee.org/document/7067404/

Ben T. Files, Vernon J. Lawhern, Anthony J. Ries, and Amar R. Marathe, (2016). **A permutation test of unbalanced paired comparisons of global field power**. *Brain Topography*, 29, 345-357.
DOI: 10.1007/s10548-016-0477-3
https://link.springer.com/article/10.1007%2Fs10548-016-0477-3

Hubert Cecotti, and Anthony J. Ries, (2016). **Best practice for single-trial detection of event-related potentials: application to brain-computer interfaces**. *International Journal of Psychophysiology* 111, January, 2017, 156-159
DOI: 10.1016/j.ijpsycho.2016.07.500
http://www.sciencedirect.com/science/article/pii/S016787601630633X

2.3.3 ICB Datasets

A total of 363 datasets were collected across 82 recording sessions, using 82 unique subjects, within the 4 specified experiments.

The total size of the EEG files is 54.12 GB, representing 94.19 hours of EEG recording.

2.3.3.1 ARL_ICB_CT2WS Dataset

A total of 147 datasets were collected across 20 recording sessions, from 20 unique subjects recruited via the GVSL.

The total size of the EEG files is 15.52 GB, representing 28.42 hours of EEG recording.

2.3.3.2 ARL_ICB_RSVP Dataset

A total of 91 datasets were collected across 14 recording sessions, from 14 unique subjects recruited via the GVSL. This total includes 88 complete datasets, and 3 abbreviated datasets.

The total size of the EEG files is 10.38 GB, representing 18.68 hours of EEG recording.

2.4 Cognition and Neuroergonomics Collaborative Technology Alliance (CaN CTA)

Difficulties experienced as US forces attempt to identify and neutralize the threats associated with the evolving security context have inspired the Army to reconsider, at a fundamental level, the capabilities and readiness of its personnel and materiel resources. Imperatives to prepare and transform the Army to meet the demands of the modern strategic environment have indeed placed science and technology in a prominent position.

Enabling technology advances are enhancing Soldier-system performance and expanding operational capabilities, however, these advances can also intensify the need to assess the Soldier's ability to perform his or her tasks under complex, dynamic, and time-pressured operational conditions. Technological advances, particularly in sensor deployment, information bandwidth, and automation, coupled with economic and political realities, will continue to place more and more responsibility on fast, distributed, and effectively independent decisions by solo or small groups of soldiers who control ever more potent defensive and offensive assets. Under such circumstances, Soldier cognitive failures in comprehension and decision-making based on an ever more complex data stream may become a critical bottleneck in Army defensive and offensive capabilities. Indeed, both failures to act and imprudent reactions can have high costs in terms of mission success, human suffering, and sociopolitical perception.

It thus becomes increasingly important that Army systems integrate knowledge of, as well as actively enhance, operator cognitive state and reactions to events. Towards this end, neuroscience-based approaches have the potential to provide revolutionary advances to foster practical solutions to address Army needs. Recent progress in the neurosciences has greatly advanced our knowledge of how brain function underlies behavior, providing our modern scientific foundations for understanding how we sense, perceive, and interact with the external world; an understanding that, if properly leveraged, can lead to improved capacity for integrating human neurocognitive function with Army system design and performance.

2.4.1 CaN CTA Program Summary

The Cognition and Neuroergonomics Collaborative Technology Alliance (the CaN CTA) program was formed by the U.S. Army Research Laboratory (ARL) as a consortium of government, industry and academic research partners to address these challenges. Here, the term "neuroergonomics" is used as originally proposed and defined by Parasuraman (2002) as "the study of the brain and body at work," and specified for military applications as "operational neuroergonomics" within the CaN CTA as building upon neuroscience, human factors, psychology, and engineering to enhance our understanding of Soldier brain function and behavior in complex operational settings, assessed outside the confines of standard research laboratories.

2.4.1.1 ARL EEG Comparison Study

Advances in neurotechnology have made it possible for researchers to investigate brain function beyond the laboratory using mobile electroencephalography (EEG) systems. Mobile EEG systems offer researchers experimental flexibility and a cheaper alternative to laboratory-based systems; however, it is unclear if their signal quality is comparable. Here we compared signals acquired from two wireless systems, Advanced Brain Monitoring (ABM) X10 and Emotiv EPOC, to signals measured from a conventional, wired BioSemi EEG system using both human participants and a surrogate phantom head. Participants performed a visual oddball task while wearing each of the three systems on different days.

2.4.1.1.1 Protocol

The research protocol information for this study is as follows:

> Organization: ARL
> Protocol Number: ARL-20098-10027
> Protocol Name: "Comparison of operationally-relevant EEG systems"
> Contract: W911NF-10-2-0022

2.4.1.1.2 Location

This study was conducted at the ARL indoor, climate-controlled Mission Impact through Neuroergonomic Design (MIND) laboratory.

2.4.1.1.3 Subjects

A total of 18 subjects participated in this study, with each subject participating in data collection on 3 separate days, wearing a different EEG system headset each day while performing a series of standard tasks.

2.4.1.1.4 Apparatus

The following equipment was allocated for data collection in this study:

> *Standard PC:* A standard computer may be used to present auditory and visual stimuli to participant, where the timing and spatial location of the stimuli are controlled by experimental presentation software, such as MATLAB or E-Prime. Participant responses may be collected using a keyboard, keypad, mouse, joystick, or microphone (verbal responses).

Simulator: The simulator is a fixed-base driver's station with three flat screen computer monitors (set at 120° from display face to display face) and a driver control system that consists of a steering wheel, foot brake, and accelerator. The driving configuration will represent the indirect driving vision system on the CAT-ATD. The computer-generated road-scene graphics are created using MINI-SIM by Real-Time technologies, Inc. The participant may be tasked to drive a simulated vehicle through a computer generated test course that requires several maneuvers, including lane changes, sharp turns, and decreasing radius turns.

Eye -tracking and Monitoring System: faceLAB4™ (Seeing Machines, Canberra, Australia) is a camera-based tracking system that allows completely non-contact operation, allowing for observation of the natural participant eye and head movement behavior at adequate spatial resolution (~0.5º). Eye and head movements, along with measurement reliability data, may be logged in real time and synchronized with the other data measures. No video record is captured by this data collection system.

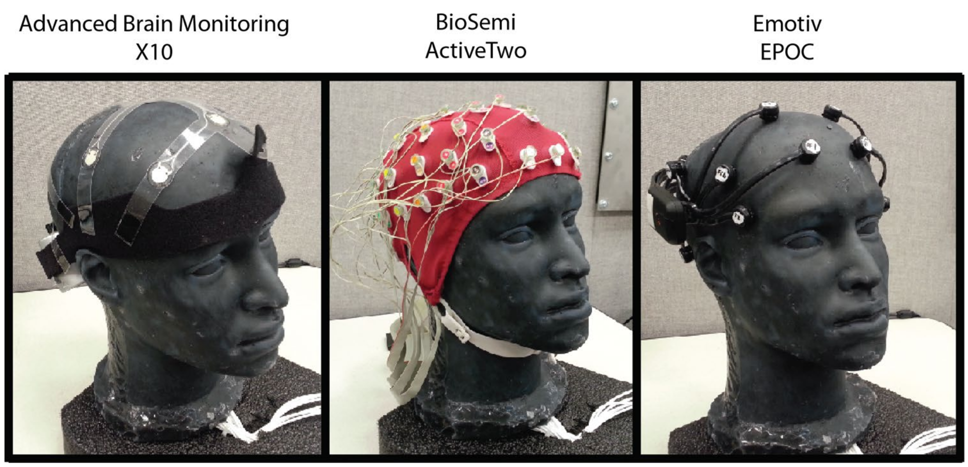


Figure 2-22. EEG Headsets and a Phantom Head

ABM EEG System: A 9-channel EEG headset (ABM, Inc., Carlsbad, CA) that transmits signals over a wireless or RS-232 wired interface. EEG recording sites conform to the international 10-20 electrode placement system. A water-soluble salinated (salty) electrode gel will be used to facilitate conductivity between the scalp and electrode surfaces and performed with strict adherence to the safety guidelines established by the Society for Psychophysiological Research (Putnam, Johnson, & Roth, 1992). The weight of this system is about twelve ounces. The ABM is shown in Figure 2-22.

BioSemi EEG System: An EEG system using an ActiveTwo amplifier and electrode cap (resembling a swim cap) with pre-amplified surface electrodes (BioSemi, Amsterdam, Netherlands), sampling at a rate of 500 Hz. EEG recording sites will be prepared in accord with the standardized international 10-20 electrode placement system (Nuwer et al., 1994) and performed with strict adherence to the safety guidelines established by the Society for Psychophysiological Research (Putnam, Johnson, & Roth, 1992). A water-soluble salinated (salty) electrode gel will be inserted into each of the electrode casings to facilitate conductivity between the scalp and electrode surfaces. Vertical (VEOG) and horizontal (HEOG) eye movements will be monitored using bipolar electrode montages attached superior and inferior to the right eye (VEOG) and both orbital fossa (HEOG). The collective weight of these electrodes is three ounces. The BioSemi is shown in Figure 2-22.

Emotiv EPOC System. A lightweight, commercially available (Emotiv, Australia) EEG headset developed for advancing BCI applications. The system consists of 16 felt pad-based electrodes within a flexible plastic frame, which comfortably fits most head sizes, and transits wirelessly to a USB-based PC receiver (no tethering necessary). Contact with scalp skin through normal hair is maintained via standard saline solution; no medical electrode gel is necessary, providing ease and comfort for users and minimal preparation. The weight of the system (dry) is 10 ounces. The Emotiv is shown in Figure 2-22.

NCTU Dry Electrode System: A small, lightweight headpiece (NCTU, Taiwan) that relays EEG data over a standard wireless data transmission protocol to a recording PC device. The system consists of NCTU's miniaturized MINDO dry-electrode sensors that work through normal hair and require no prior skin preparation, nor the use of conductive gels to facilitate electrical contact with the scalp. The headpiece positions the sensors at a sampling of the positions of the standardized international 10-20 electrode placement system. The weight of this system is twenty-five ounces.

QUASAR EEG System: A lightweight, user-adjustable headpiece (QUASAR, Inc., San Diego, CA) that relays EEG data to an external Base Station wirelessly or via a RS-232 wired interface. The system consists of QUASAR's hybrid, high impedance, dry-electrode sensors that work through normal hair and require no prior skin preparation, nor the use of conductive gels to facilitate electrical contact with the scalp. The sensors can be individually adjusted to improve their contact with the scalp. The headpiece positions the sensors at the nominal F3, Fz, F4, C3, Cz, C4, P3 and Pz positions of the standardized international 10-20 electrode placement system. An additional reference sensor is placed at the nominal P4 position. The weight of this system is ten ounces.

2.4.1.1.5 Demographics

No demographic data is provided with this dataset.

2.4.1.1.6 EEG Comparison Study Tasks

A number of tasks were defined for subjects to perform while instrumented with each of the EEG systems.

***Note**: The only data in the repository for this study is the combination of the VEP/Oddball task performed while subjects were instrumented with the BioSemi EEG system.*

2.4.1.1.6.1 Visually-Evoked Potential/Visual Oddball (VEP/Oddball)

(Approx. 4 minutes)

Participants will sit comfortably in front of a computer monitor. To begin, participants will fixate on a black fixation cross, centered on the screen. On each trial, a visual stimulus will appear in the center of the screen for 150 milliseconds followed by 1000-1200 milliseconds inter-stimulus interval containing only the fixation cross. On 12% of the trials, the stimulus will be a picture of an insurgent, and on the other 88%, the stimulus will be a picture of a US Soldier. Participants will be instructed to press a response key after each stimulus. One response key will be designated for the insurgent and a different response key for the US Soldier. The stimulus-response mapping will be counterbalanced across participants. This procedure modifies previous oddball research (Polich & Kok, 1995; Kerick et al, 2009) with operationally relevant images.

2.4.1.1.6.2 Resting State

(Approx. 16 minutes)

Participants will sit comfortably in a chair. There will be twelve 1-minute resting baselines, rotating through three conditions: resting with eyes closed in a well-lit room (total of 4), resting with eyes open in a well-lit room (4), and resting with eyes open in a dark room (4), with 45 seconds in between each 1-minute session to allow the participant to change states. Participants will be instructed to try to be as restful as possible. An auditory cue on the computer will signal when the participant is to switch states. This procedure modifies previous research (Tomarken et al, 1992).

2.4.1.1.6.3 Error-Related Potential

(Approx. 16 minutes)

Participants will sit comfortably in front of a computer monitor. To begin, participants will fixate on a black fixation cross-centered on the screen. On each trial, a visual stimulus will appear in the center of the screen and remain until the participant responds. Each trial requires the participant to press a button to move non-green box to overlap the green box, and the box moves 1 second later. On ~20% of trials, the box moves in the opposite direction of the button press.

2.4.1.1.6.4 N-back

(Approx. 11 minutes)

Participants will sit comfortably in front of a computer monitor. To begin, participants will fixate on a black fixation cross-centered on the screen. On each trial, a single letter of the alphabet will appear in the center of the screen for 300 msec followed by an inter-stimulus interval containing only the fixation cross. For a block of 30 trials, the participant will respond if the current letter matches the letter seen on the last trial (1-back). For the next block of 30 trials, the participant will respond if the current letter matches the letter on the screen two trials ago (2-back). The blocks will continue to alternative between 1-back and 2-back, with the order of the blocks counterbalanced between participants.

2.4.1.1.6.5 Artifact Susceptibility

(Approx. 15 minutes)

Participants will initially sit comfortably in a chair. There will be 7 seated movement conditions, lasting approximately 1 minute each with 20 seconds in between tasks: vertical jaw movement, eyes blinks, lateral eye movements, vertical eye movements, raising the eyebrows, head rotations side to side, shoulder shrugs, torso rotations from the hips from side to side. The participant will be cued what direction to move and when to move by an auditory stimulus. Following the seated conditions and a 1-minute break, the participant will stand up and sit down thirty times and then remain standing to march in place for 1 minute. They will also walk at a relaxed pace for 5 minutes. Finally, the participant will sit comfortably in a chair with their eye closed while the experimenter tests two sources of electrical noise: turning on and off the lights and turning on and off a two-way radio. This procedure extends prior research (Estepp et al, 2009; Kerick et al, 2009).

2.4.1.1.7 POC

Dave Hairston (william.d.hairston4.civ@army.mil), ARL, Research Collaborator
Jean Vettel (jean.m.vettel.civ@army.mil), ARL, Research Collaborator

2.4.1.2 DCS Finger Tapping Experiment (DCS_FT)

When recorded at the scalp using standard electroencephalography (EEG) techniques many neural signals exhibit poor signal-to-noise ratios (SNR). This often makes robust analysis difficult, if not impossible. The purpose of this study is to acquire a data set that enables the investigation, development, and validation of novel methodologies for better detection and analysis of scalp-based neural signals. The signals of interest for this study are the phase-locked responses (also known as motor potentials) and power spectral changes related to deliberate self-paced finger movements. We chose these signals because there has been a considerable amount of prior research into the cortical dynamics of finger movements. Also we feel that no one has yet reliably shown the capacity for differentiating brain activity associated with individual finger movement using scalp level electrical data. We feel that by working towards this goal we will greatly enhance our current understanding of the type of data EEG provides as well as demonstrate the utility of EEG for detecting previously disregarded, subtle cortical changes.

2.4.1.2.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL/DCS
Protocol Number: ARL-113-069
Protocol Name: "Investigating Cortical Components of Deliberate Finger Movements using High Density (256 Channel) Electroencephalography"
Contract: W911NF-10-2-0022

2.4.1.2.2 Location

This study was conducted at ARL's MIND lab at Aberdeen Proving Ground, MD.

2.4.1.2.3 Subjects

A total of 13 subjects were recruited for this study, with each subject participating in a single session and performing the task one time, producing 13 data recordings.

2.4.1.2.4 Apparatus

This project required the use of two standard PCs. One PC was used for EEG data collection and one was used for stimulus presentation. Stimuli were presented with custom software, developed using the E-Prime software package produced by Psychology Software Tools, Inc., indicating which finger should be used for each task. One additional monitor and response pad were provided for subject input. The response pad contained five buttons, four of which were used in this task.

Subjects were instrumented with a BioSemi 256 (+8) EEG channel system with 4 eye and 2 mastoid channels recorded, with a sampling rate of 1024 Hz.

2.4.1.2.5 Demographics

No demographic data is provided with this dataset.

2.4.1.2.6 Finger Tapping Tasks

There was one task specified for this experiment, which is described as follows:

FT:
Duration: ~80 minutes

Task requirements: Subjects were prompted to tap a button on a response pad repeatedly with each of 4 different fingers. The sequence was as follows:

Repeat the following "block" of subtasks 8 times:
<begin block>
Left hand, middle finger (2 minutes)
<end block>
<begin block>
Left hand, index finger (2 minutes)
<end block>
<begin block>
Right hand, index finger (2 minutes)
<end block>
<begin block>
Right hand, middle finger (2 minutes)
<end block>
Break (self-paced)

Perform the following block of subtasks 1 time:
<begin block>
Randomly tap fingers of subjects choosing (4 minutes)
Note: maintain the pace of the previous tasks
<end block>

Each finger was to be tapped at a consistent pace, while trying to maintain a 4-5 second separation. Subjects were encouraged, though not required, to avoid mental counting in favor of simply developing and adhering to an internal rhythm. This was to minimize contamination from any secondary cognitive tasks.

Sessions: Each subject was scheduled to perform the Finger Tapping task 1 time.

2.4.1.2.7 POC

Sandhya Rawal, Primary Investigator
Kaleb McDowell (kaleb.g.mcdowell.civ@army.mil), Associate Investigator
Stephen Gordon (sgordon@dcscorp.com), Associate Investigator
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer

2.4.1.3 DCS CANCTA Operator Dynamics of Event Appraisal Experiment (DCS_CANCTA_ODE)

In response to current and projected needs associated with the optimization of Warfighter interactions with typically information-dense and highly dynamic, complex computerized interfaces, the CaN CTA has proposed to explore how research in cognitive neuroscience could best be used to design individualized real-time neuroergonomic systems aimed at improving situational awareness and decision-making under stress. To this end, an integrative research program has been described that would extend the range and depth of known principles of and established methods for estimation of neurocognitive state, event appraisal, and behavioral intentions of Army operators of such systems.

To date, efforts within the CaN CTA have been advancing objectives largely associated with methods for neurocognitive state estimation and inferences of behavioral intent. Although considerable progress has been made in these directions, methods for assessing operator appraisal of the operational context, as well as the efficacy of actions within it, are only beginning to be examined within the CTA. Indeed, the availability of methods and tools for extracting comprehensive information regarding how the operator is reacting to, planning actions in, and appraising effects of those actions on the unfolding task environment will be an essential component of advanced computational characterizations of individual differences in neurocognitive performance. Moreover, such tools will broaden our insight regarding when and how to most appropriately intervene in the ongoing performance of an operational task.

To meet the objectives just discussed, significant continuing efforts within the CAN CTA involve the application of computational and machine-based operator state classification and 108 prediction algorithms to data derived from increasingly advanced sensor systems. To date, however, the information driving development of these algorithms has largely, though not exclusively, been derived from electroencephalographic (EEG) data. Certainly, the intensive focus on the information contained in complex EEG signals is critical to addressing the scientific barriers outlined earlier in this document (section 2.2). Yet, it is equally crucial to recognize and address the fact that, in the absence of information from other sources that naturally co-vary with EEG – such as autonomic function and overt motor behavior – an understanding of the true nature of the relations between neural dynamics, operator state, and operational behavior will remain elusive.

Indeed, given the spatial limitations and constrained set of behavioral options for in-vehicle (mounted) operations, as well as the requisite computational overhead, there is a question as to the practical value of recording a dense array of kinematic variables describing overt behavior. Yet, as has been argued earlier in this document (see section 2.2, B3), information regarding the whole of the objective behaviors enacted by the Soldiers in the operational context is essential for constructing an understanding of the neural patterns observed in EEG and thus for drawing inferences regarding the perceptual-cognitive processes being engaged during military tasks.

We believe that even within such a constrained setting as a crewstation there is information that can be leveraged beyond EEG signals to improve the classification process and that such information may be useful to help identify states with unique, yet consistent, brain dynamics. For example, through the use of computer-vision technologies that register graph topologies with individual facial features (see Figure ODE 1) as well as intelligently selected and placed sensors we intend to capture and record continuous streams of data related to individual motor (facial, ocular, upper limb) behavior, coupled with EEG and other surface physiological data (heart rate, EMG), while participants perform militarily-relevant tasks. This also moves towards more completely implementing the consortium's notion of "Mobile Brain/Body Imaging" (MoBI) (Makeig et al, 2009). Though, while current MoBI concepts are intended to capture gross body movements in less-constrained environments we will focus on the types of behavior most applicable to and readily retrievable in the highly-constrained crewstation environments that are typical of armor-laden military vehicles.

The overarching goal of this research is to develop and validate methods for enabling and making inferences about operator event appraisal processes as reflected in changes observed in cognitive and emotional state variables during the execution of tasks in operational environments. The aims underlying this goal include:

<u>Aim #1</u> is to develop and experimentally validate an integrated system capable of recording and synchronizing high-density, multimodal data on human physiology and behavior. To enable system validation, we will record and explore the patterns of correlation and co-variance among a variety of psycho-physiological and behavioral response variables. Measures will be derived from EEG, EOG, EMG, ECG, limb, head, and gaze position tracking, facial expressions, verbalizations, salivary samples for determination of concentrations of cortisol and testosterone, and the participant responses to the tasks using the mouse, keyboard, and/or response pad. Subjective state and changes thereof will be probed by means of standard questionnaires, such as the Positive and Negative Affect Scale (PANAS) and the Self-Assessment Manikin (SAM), to validate state changes detected in the behavioral and physiological data.

Initially, the primary efforts undertaken to support this aim focus on the integration of the previously-discussed multimodal data using physiological recording systems already available at DCS and through collaborators at the Translational Neuroscience Branch at ARL. Along with

EEG, the additional data modalities will be recorded and, of course, engineers will subsequently test and verify the sensor system capabilities regarding the sampling, synchronization, and logging of the multimodal data. Later efforts will use virtual environment simulation created with SCCN's SNAP tools and a crewstation interface that will be a designed to emulate key aspects of TARDEC's current Warfighter Machine Interface (WMI). In this latter effort, we intend to replicate, in improved form, simulation environments and experimental systems and protocols already developed to allow participants to drive (without concomitant motion simulation) around a virtual urban environment emulating a typical military setting (e.g. a generic village located in the Middle East) while tasked with standard target detection and communications tasks.

Aim #2 is to develop computational strategies that will facilitate sensor management in order to enhance the acquisition and processing of multimodal data for studies of human neurocognitive performance in operational environments. Our approach will involve applying an intentionally-selected set of algorithms to human physiologic and behavioral data which, when performing reliably, will enhance the online acquisition of complex data sets as well as their later offline processing. We will also apply several machine learning algorithms to characterize individual data streams with respect to various data quality measures in order to establish their reliability and robustness.

In support of this aim, we will apply a variety of machine learning algorithms to the integrated and synchronized data set that will be defined around the use of features extracted from the additional modalities of data with the goal to create a set of filters that will serve to automate the artifact rejection process. We will then focus on those modalities that are expected to have strong correlation, e.g. 1) video-based eye tracking and EOG, 2) facial expressions, facial EMG, and frontal/temporal EEG, 3) head movements and upper shoulder/neck EMG, and 4) upper shoulder/neck EMG and crewstation interactions, to name a few. Using standard techniques, we intend to develop specific data transformations to maximize correlation across modalities. The transformed data will then be fed into machine learning algorithms that either look for specific fault conditions, or use Bayesian inferencing to determine the most likely state (i.e. "fault" or "not fault").

Finally, when applicable, and possible, we will attempt to supplement faulty data with data derived from a secondary source. For this to be maximally successful, it will require an understanding of the intended use of the faulty data, as well as the sensitivity of the user to changes in that data. At present, we intend simply to provide supplemental data along with an indicator signal that the data source has changed. Finally, while we recognize that sensor management is typically viewed as an online tool, we plan to develop our approaches first using post-processed data and then port these, as time and resources allow, to online conditions.

One future direction for this work would be to perform a sensitivity analysis of processing methods that use such data, e.g. machine-learning classifiers and state estimators, in order to determine which features are most critical for robust performance, how does performance degrade with feature degradation, and how does performance differ if a secondary data source is used?

The third aim of this effort is to use EEG and additional modalities of data to enable the classification of operator appraisal processes while they perform both simple and complex cognitive-motor tasks.

For this aim, we consider appraisal processes to be reflected in two ways. The first builds from the theoretical work of Scherer (c.f. Scherer, Schorr, & Johnstone, 2001), in which we will consider appraisal to be a fundamental component of affective processing. As such, we expect appraisal processes to be reflected in assessments of affective state (i.e. frustration, contentment). The second type of appraisal is assumed to be related to processing of the performance of the task at hand. That is, as a task unfolds, we expect the operator's assessment of their own performance to be reflected in instantaneous reactions following the selection and execution of an action. For example, in a difficult threat detection task, pilot testing has suggested that we might see changes in facial expressions when an operator immediately detects that he or she has wrongly identified the type of target that they just viewed. Thus we may see stereotyped responses in situations provoking errors (e.g. a grimace indicating dissatisfaction) versus correct responses (i.e. a smile or a nod indicating satisfaction), which will likely be classifiable based on differential patterns of EMG, facial expression, and, perhaps, autonomic variables such as heart rate.

2.4.1.3.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL/DCS
Protocol Number: ARL 13-043
Protocol Name: Operator Dynamics of Event Appraisal I
Contract: W911NF-10-2-0022

2.4.1.3.2 Location

This study was conducted at ARL's MIND lab at Aberdeen Proving Ground, MD.

2.4.1.3.3 Subjects

A total of 17 subjects were recruited for this study, with each subject participating in a single session and performing the task 4 times, producing 67 usable data recordings (and one that was not usable).

2.4.1.3.4 Apparatus

This project required the use of four computing platforms (tower systems or laptops) operating in a distributed architecture, as illustrated in Figure 2-23. The purpose for each system is as follows:

1. Experimental Control system
    a. Simulation and data recording control interface
    b. Provides stimuli to subject monitor and to speakers
    c. Provides common event code / triggers for all data streams
2. Data logging system 1
    a. Operator inputs (keyboard, mouse)
    b. Webcam audio feed
3. Data logging system 2
    a. BioSemi data logging (EEG, EOG, ECG, EMG, EDA)
    b. Trigno system logging (accelerometer)
4. FaceLAB system
    a. Eye-tracking system processing and data logging

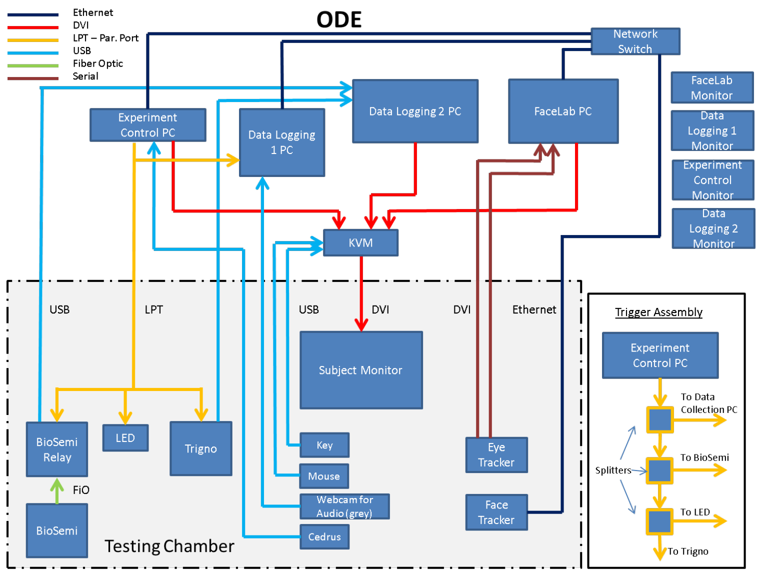


Figure 2-23. ODE System Architecture

Subjects were instrumented with a BioSemi 64 (+8) EEG channel system with a sampling rate of 1024 Hz. External channels connected to the BioSemi DAQ unit were used as follows:

EX1: Left mastoid
EX2: Right mastoid
EX3: Right eye horizontal EOG
EX4: Right eye vertical EOG
EX5: Forehead EMG (align w/Fz)
EX6: Left jaw EMG (lower, centered about masseter)
EX7: Left jaw EMG (upper, centered about masseter)
EX8: Chest ECG

Subject also wore a Trigno system, for the purposes of capturing Accelerometry data via sensors at the following locations:

1: Right neck
2: Left neck, center of sternocleidomastoid
3: Left trapezius
4: Right trapezius
5: Right wrist
6: Left wrist
7: top center of EEG cap
8: Used to receive event trigger codes from experimental control system

Additionally, a FaceLAB eye-tracking system was used with cameras positioned about the subject monitor at the test station, and a webcam was used to record audio during the session.

2.4.1.3.5 Demographics

No demographic data is provided with this dataset.

2.4.1.3.6 ODE Tasks

There was one task specified for this experiment, performed in 4 parts, which are described as follows:

ODE:
Duration: ~80 minutes

Run 1: PRACTB – Calibration Run

Subjects performed a variety of artifact inducing and baseline calibration tasks (see Figure 2-24) including:

2 minutes of fixating on a cross hair
2 minutes of visually tracking a ball as it bounced around the screen
Speaking 20 short phrases out loud
Making a happy, sad, or neutral face (6x each)
20x eye blink
20x jaw clench
20x eye brow raise
35x saccadic eye movements to a random point (points were distributed as they are for the 5 on a dice)
2 minutes of watching a video of the ODE scene with no targets
20x reaction time measures (push a button as soon as you seen a stimulus). 20x right hand, 20x left hand.

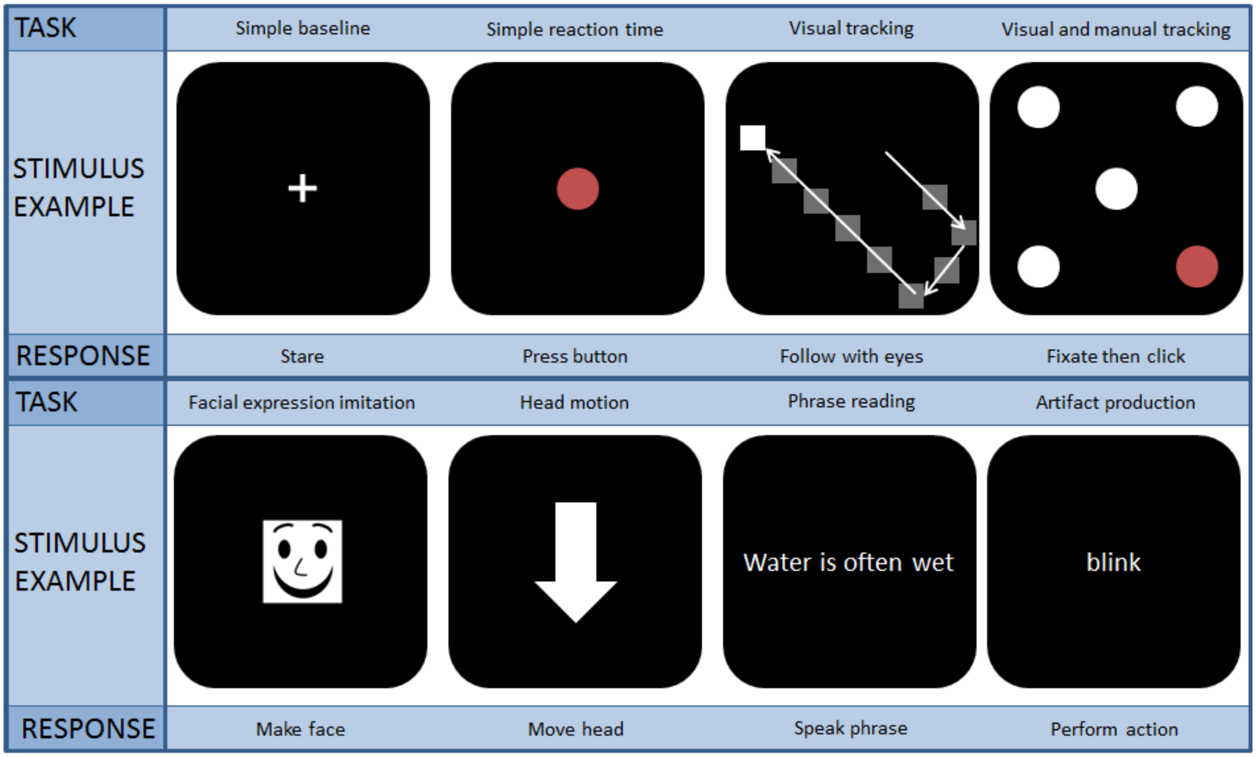


Figure 2-24. ODE PRACTB Task

Game Run 1:

Subjects were instructed to watch for the appearance of visual targets (approximately 1 every 3 secs) and to discriminate and report the type of target (left or right hand button press) as quickly and accurately as possible. The process is shown in Figure 2-25. Targets consisted of humans with weapons, humans without weapons, tables with a tablecloth (or sheet) covering what's underneath, and tables without a cloth. Points were given to the subject (visual feedback after every target) based on speed and accuracy. Subjects were driven through a city during this task. At different times in the video, a dense fog was overlaid on the scene.

After the end of 15 minutes, subjects watched the same 2-minute video as shown in the PRACTB

Game Run 2:

Same as game run 1, but in a different part of the city.

After the end of 15 minutes, subjects watched the same 2-minute video as shown in the PRACTB

PRACTB2 – Post-run calibration

Subjects performed a reduced version of the same artifact inducing tasks and calibration tasks as in the PRACTB.

2 minutes of visually tracking a ball as it bounced around the screen
Making a happy, sad, or neutral face (6x each)
10x eye blink
10x jaw clench
10x eye brow raise
35x saccadic eye movements to a random point (points were distributed as they are for the 5 on a dice)

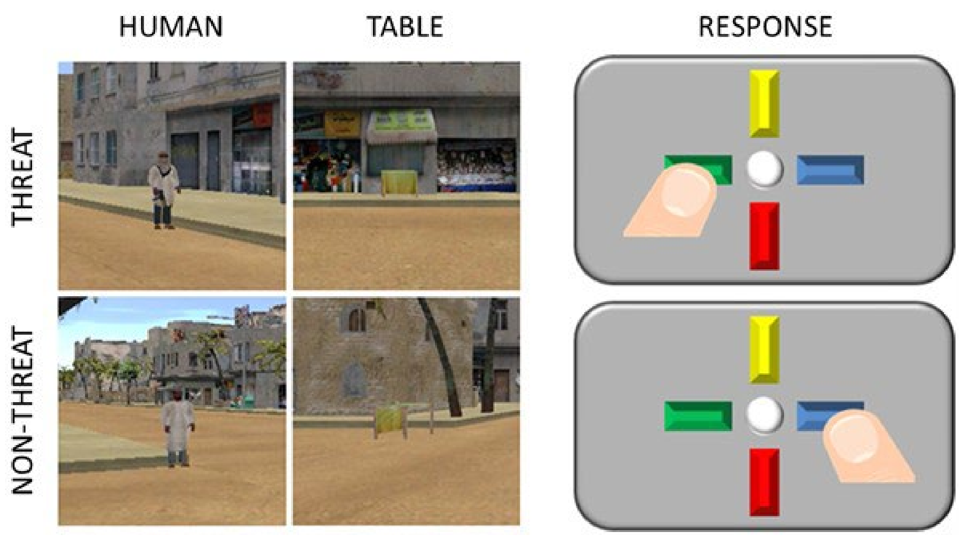


Figure 2-25. ODE Target Detection and Classification Task

Sessions: Each subject was scheduled to perform the 4 parts of the ODE task 1 time.

2.4.1.3.7 POC

Kelvin Oie, (kelvin.s.oie.civ@army.mil), Primary Investigator
Jason Metcalfe (jason.s.metcalfe2.civ@army.mil), Associate Investigator
Antony Passaro (antony.d.passaro.civ@army.mil), Associate Investigator
Keith Whitaker (keith.w.whitaker1.civ@army.mil), Associate Investigator
Stephen Gordon (sgordon@dcscorp.com), Associate Investigator
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer

2.4.1.4 NCTU Real-World Neuroimaging Vehicles Driving Environments Experiment (NCTU_CANCTA_RWN_VDE)

Soldiers performing sustained military operations often function for extended periods in stressful environments with fractionated or no sleep. It is well-established that fatigue, whether due to acute or chronic sleep deprivation, extended time-on-task or the interaction between sleep- and task-related factors, is associated with neurocognitive performance decrements across a broad range of perceptual, cognitive and motor functions (Killgore, 2010; Lim & Dinges, 2010). Motor vehicle crashes account for nearly one-third of U.S. military fatalities annually and are the leading cause of US military fatalities (Krahl et al., 2010). Further, one of the leading causes of vehicle accidents is driver fatigue (NHTSA, 2011). Fatigue, as well as stress, has been shown to dysregulate executive attentional control mechanisms underlying performance (Bishop, 2008; van der Linden et al., 2003).

Although much research has been devoted to understanding relations between brain activity and fatigue states of drivers, the vast majority of this research has been conducted in driving simulators (SIM) under highly controlled laboratory conditions so it's not known how well findings generalize to complex real-world (RW) driving. One issue with investigating fatigue in the laboratory is the artificial manipulation of sleep deprivation. Most researchers have employed full or partial sleep deprivation paradigms. In full sleep deprivation paradigms, subjects are denied sleep continuously over a 24-hr period or longer, whereas in partial sleep deprivation paradigms subjects are restricted to just a few hours of sleep (e.g., 2-6 hours) over a period of 3-4 days. In the real world, full sleep deprivation is a much less common than partial sleep deprivation (Durmer et al., 2005). However, even partial sleep deprivation paradigms require control over the sleep-restricted periods in a regimented manner, which do not accurately reflect sleep patterns of individuals in the real world. Therefore, we propose an alternative paradigm for investigating the effects of real-world fatigue on performance in both SIM and RW driving experiments. Specifically, we will leverage a Daily Sampling System (DSS) developed in Program Year 4 to monitor and track subjects' daily variations in sleep patterns and perceived levels of stress and fatigue as experienced by subjects naturalistically on an everyday basis. The DSS will automatically evaluate each subject's daily levels of fatigue based on actigraphy, sleep diaries, and subjective reports and schedule subjects for experiments along a continuum of levels of fatigue. For the purposes of this research, note that while sleepiness may be considered an important component of fatigue, the terms are not synonymous.

Another issue with investigating fatigue in the laboratory is the artificial driving environment inherent in driving simulators. Realistic driving conditions are difficult to simulate because there is no element of danger or real consequences for degraded driving performance in SIM driving, as is evident in RW driving. In order to overcome this issue, we have planned a series of SIM experiments designed to simulate increasingly more complex, realistic driving environments in a ride motion simulator, but under experimental control, while also planning an observational study in which we will examine data from subjects driving in the real world.

By addressing these two issues, we will be able to better understand how the brain functions during real driving under the demands of real fatigue. EEG, eye tracking, driving performance, and subjective report data will be recorded, integrated, and analyzed from a large number of subjects in both SIM and RW driving environments over repeated sessions across different driving conditions. Experiments comprising this 3-yr effort will generate extensively large and complex data that will also be leveraged to generate a unique and unprecedented database in terms of the number and diversity of experiments, subjects, measures and meta-data (i.e., "Big Data"). The database will facilitate hypothesis-driven research focusing on brain-behavior relations in real-world environments as well as data mining and exploratory research. The present research plan proposes analysis of within- and between-subjects differences, analysis of SIM versus RW differences, comparisons of approaches in signal processing, statistics and multifactorial analyses, data integration/fusion, feature extraction, data reduction, collaborative filtering and clustering, and modeling and prediction algorithms.

The objective goal for understanding real-world fatigue within vehicle driving environments is:

> **Objective Goal**: Image and interpret real-world fatigue during real driving with a quality that enhances the fundamental understanding of the underlying brain processes.
> The challenges encountered in this three-year study will be many-fold, as will the scientific, engineering, and analytical efforts aimed at meeting those challenges. In attempting to understand brain activity during real driving, it will be essential to extend driver fatigue monitoring from small-scale laboratory experiments to practical real-world applications. Driving provides a constrained trajectory that allows identification among activities in a simulator and those in the real world. However, the complexity of the decision space and the diversity of responses within and between subjects require a substantial data-collection and analysis effort to validate any proposed monitoring tools. Therefore, an experimental framework that leverages both the experimental control offered in the simulation environment (e.g., where the subjects' driving performance can be measured against their response to experimentally induced vehicle perturbations), and the real-world aspect of driving in everyday environments, is expected to produce a useful aggregation of data that can be analyzed using a common approach. However, in the event potential difficulties are realized in collecting the real world driving data (e.g., owing to IRB concerns), a minimally sufficient goal for the project has also been established:
>
> **Threshold Goal**: Image and interpret real-world fatigue during simulated driving within realistic virtual environments.
> For each version of the research goal, the key concept is *real-world fatigue as it pertains to a driving task*. Use of the DSS to ensure subjects are in the intended target state during experimentation is envisioned as an effective and innovative technique for qualifying the data collection efforts within this project. It is expected to support the study of

generalizable driving fatigue models by exploring the brain dynamics associated with different fatigue levels. Such models will rely on foundational EEG analysis to assess the applicability and reliability of different neuro-markers in signaling fatigue, as well as a multi-aspect analysis approach for leveraging additional sources of physiological instrumentation data, and subject behavioral data.

The overall research plan for understanding real-world fatigue within vehicle driving environments entails a series of longitudinal studies that employ a typical three-phased approach involving experimentation to collect driving data, followed by data management processing to correlate physiological measurements with environmental events via a standard methodology, and data analysis to interpret a driver's cognitive state based on context. Separately, the studies will yield multi-aspect data sets that feature physiological measurements of subjects performing a driving task in either a simulated or real-world environment. In the aggregate, the studies will produce a substantial amount of driving data that spans environmental conditions and subject states of fatigue and stress. Thus, this three-year project will require careful characterization of the data and metadata so that comparisons can be accurately and effectively made under a variety of simulated and real-world conditions. Additionally, the data must be prepared for analysis using tools and techniques that are suitable for working with expansive data sets.

Note: The design of the real-world driving study referenced above was modified in PY7-8. Instead of studying driver fatigue, the researchers added a participant, and attempted to track driver-passenger communication as a function of emotional valence (manipulated by use of humor), and determine whether or how these factors influence the development of passenger trust in the driver. The study is titled "Assessment of Intra- and Interpersonal Brain and Behavioral Dynamics of Driver-Passenger Dyads during Real-World Driving", and assigned the project identifier RWN-VDEDP (Real-World Neuroimaging in Vehicle Driving Environments with Driver and Passenger). This study is complete at the time of this writing, and data will be added to SANDR after curation.

2.4.1.4.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL/NCTU
Protocol Number: ARL 14-088
Protocol Name: "Simulated Driving under Conditions of Real-World Fatigue and Stress"
Contract: W911NF-10-2-0022

2.4.1.4.2 Location

This study was conducted at the National Chao Tung University (Taiwan) using the motion vehicle simulator (MVS).

2.4.1.4.3 Subjects

A total of 17 subjects participated in this study, with each subject participating in data collection on 3 separate days, performing multiple tasks per day, producing 855 data recordings.

2.4.1.4.4 Apparatus

This project required the use of a subject measurement platform and several instrumentation devices, described as follows:

Driving Simulator. The virtual-reality-based highway driving experiments will be conducted in a driving simulator consisting of a real vehicle mounted on the 6-DOF motion platform in a sound-reduced room. The driving simulator mimics realistic driving situations. All the control and stimuli system are developed by the C++ software.

Electroencephalography (EEG). A 70-Channel EEG amplifier system (SynAmps^2, Compumedics Inc.), consisting of 64 monopolar, 4 bipolar and 2 high-level channels will be used to record brain electrical activity during eight different experimental recording sessions. Each channel has a dedicated 24-bit A-to-D converter, to ensure the most accurate sampling available.

Eye Tracking System. An SMI RED eye tracking system (SensoMotoric Instruments) will be used for measuring eye positions and eye movement during eight different experimental recording sessions.

Actigraphy. Wrist-worn actigraphs (Fatigue Science ReadiBand) will be provided to all participants to enable objective and accurate characterization of their sleep timing, duration, and quality, as well as model-based estimates of performance effectiveness, on a daily basis over the duration of the study. The ReadiBand actigraphs are small lightweight computerized accelerometer-based devices that digitize movement in six dimensions (x, y, z, yaw, pitch, and roll). Data from the actigraph is stored locally on the device and is retrieved by uploading the data to a computer via USB antenna and the data can be uploaded either to a cloud server anonymously for data archiving and sharing with collaborators or to the local hard drive of a computer. If the cloud server is used, no personally identifiable information (PII) will be associated with the data.

2.4.1.4.5 Demographics

No demographic data is provided with this dataset.

2.4.1.4.6 RWN_VDE Tasks

Participants performed a series of tasks within this study, as shown in Figure 2-26.

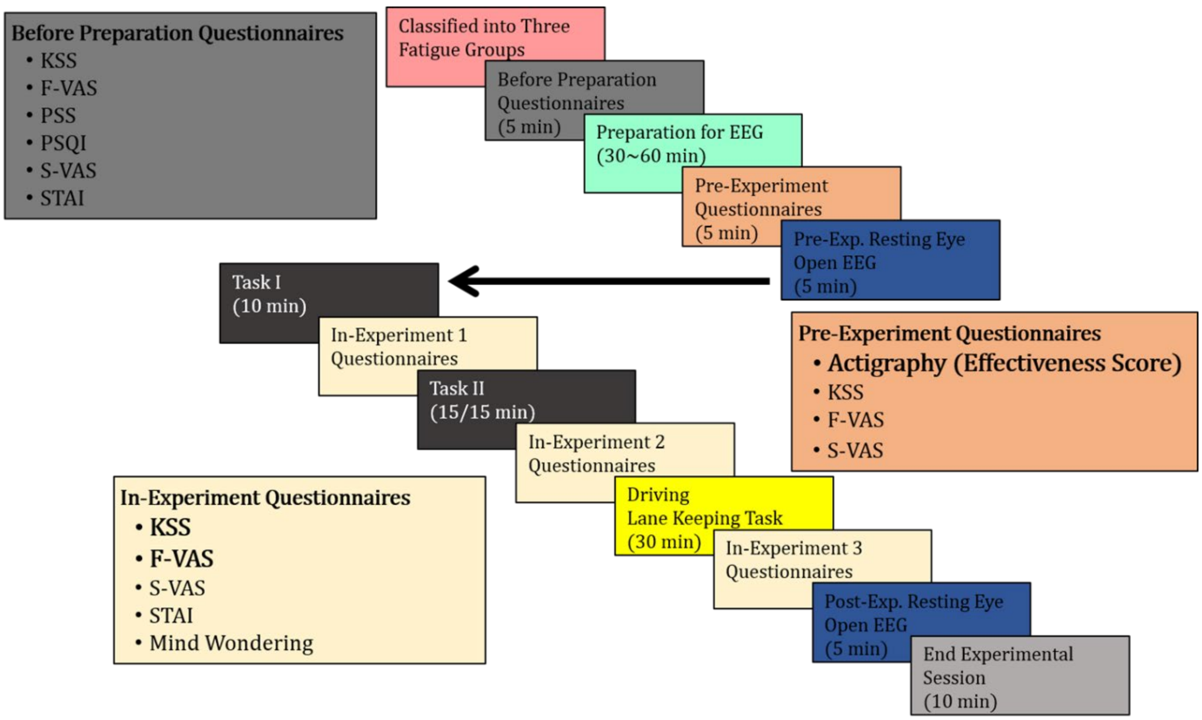


Figure 2-26. RWN-VDE Experimental Procedures

2.4.1.4.6.1 Psychomotor Vigilance Test (PVT)

Each participant sat in front of a desktop computer and completed ten minutes of the PVT paradigm. Subject reaction time was measured according to a button press after the appearance of a red dot in the center of the screen.

A response was regarded as 'valid' if the reaction time (RT) was between 100 ms and 1.2 seconds; otherwise, the response was recorded as a lapse. For performance calculation, an RT of less than 500 ms was accepted as a 'correct' response.

2.4.1.4.6.2 Lane-Keeping Task

To implement the driving task (see Figure 2-27), the virtual reality scene was constructed with the World Tool Kit (WTK) program, a C-based 3D graphic library. The program simulated driving a car at a certain speed (100 km/hr.) on the highway at night and automatically drifted away from the cruising lane to the left or right side with equal probability. Participants had been instructed to steer the vehicle back to the cruising lane as fast as possible after becoming aware of the deviation. If the participants did not respond to the lane-perturbation event, falling asleep for example, and then the vehicle could hit the left or right curb of the roadside within 2.5s and 1.5s, respectively. The vehicle would then continue to move along the curb until it returned to the original lane.

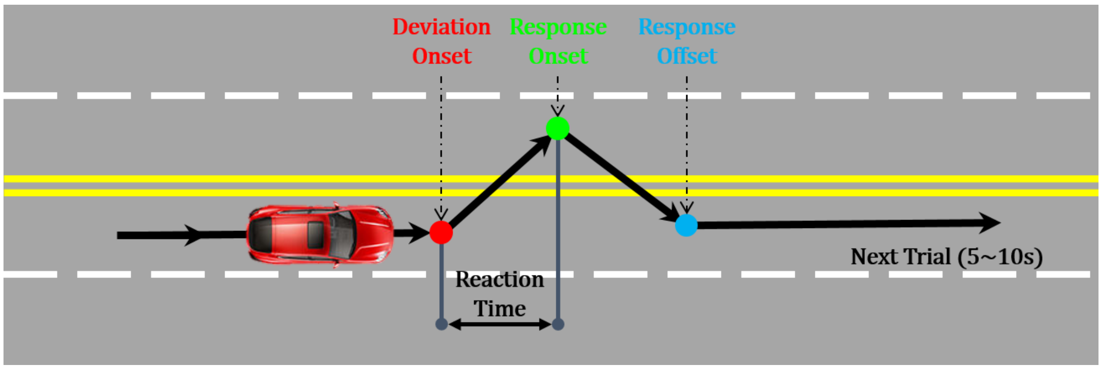


Figure 2-27. A Bird's Eye View of the Event-Related Lane-Departure Paradigm

Each lane-departure event is defined as a "trial" which includes three critical moments: "deviation onset" is the moment when the car starts to drift away, "response onset" represents the moment when the participant perceives the drift and begins to steer the car back to the cruising lane, and "response offset" is the moment when the car returns to the center of the cruising lane, and the participant ceases to rotate the steering wheel. The next lane-departure event occurs again 8 to 10 sec after the "response offset." The reaction time is defined as the interval between deviation onset and response onset in a trial. There were no other vehicles or stimuli that might disturb the driver's attention. This was intentional, in order to create a driving condition likely to induce fatigue. Participants' cognitive states and driving performance were monitored via a surveillance video camera and the vehicle trajectory throughout the experiment.

2.4.1.4.6.3 Dynamic Attention Shifting (DAS)

In the study, we proposed the two types of experimental tasks, including a lane-keeping driving task (LKT) and dynamic attention shifting task (DAS). In the lane-keeping driving task, subjects were seated in a vehicle and driving scenes were simulated at high speed (100 KMH) along a virtual 4-lane road (two lanes in each direction) without other traffic. Throughout the experiment, the computer program generated a random perturbation (deviation onset) and the vehicle simulator drifted away the cruising lane to left or right lands with equal probability automatically. At lane-departure events onset, subjects were required to steer the vehicle back to the cruising lane as soon as possible using the steering wheel (response onset), and hold on the wheel after the car returned to the approximate center of the cruising lane (response offset). The subject's reaction time (RT) to each lane departure trial is defined as the interval between deviation onset and response onset.

At the deviation onset, the car was randomly drifted either to the left or right during the lane-keeping driving task. When the subjects detected the deviation, they were instructed to steer the car back to the cruising lane quickly by turning the steering wheel. The latency between the deviation onset and turning the steering wheel was defined as the reaction time (RT), as shown in Figure 2-28.

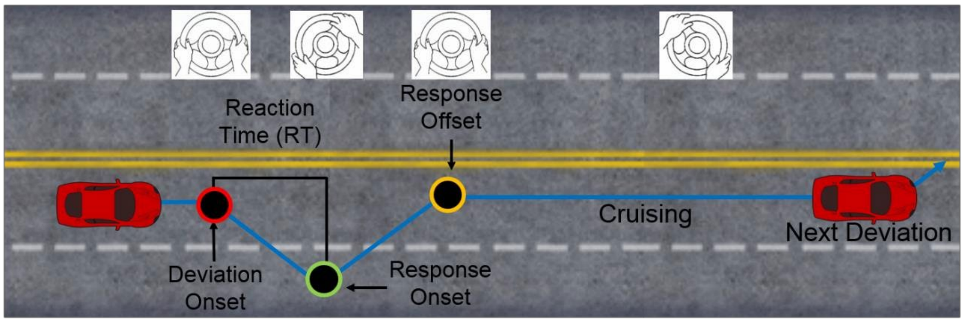


Figure 2-28. Reaction Time and Response Time Metrics

In the dynamic attention shifting task there are multiple forms of stimuli, including spoken words and written words. Subjects were asked to differentiate whether the stimulus presented in either visual or auditory were a given target or not, which mimicked driving attention shift event. At the beginning of each "round" a warning cue, the word "attention", was played on the left (right) screens or by the left (right) speakers was followed by a series of stimulus. Each round contained 4-6 stimuli and subjects had to respond to the targets and ignore the non-targets. The auditory stimuli were two-tone patterns. An example of a visual stimulus and response (see Figure 2-29. Event-Related Target Identification Paradigm) shows that the visual stimuli were presented as red letters.

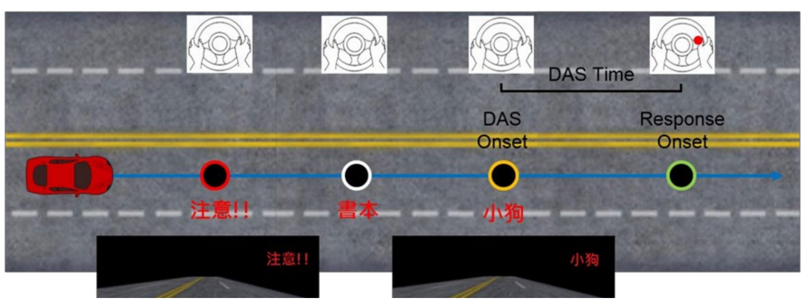


Figure 2-29. Event-Related Target Identification Paradigm

Among the words, subjects were asked to respond to animal words and to ignore the non-animal words. When the subjects detected that stimuli on the left (right) screen or from the left (right) speaker were animal types, and matched the type of warning, they were asked to press the left (right) button mounted on the steering wheel (see Figure 2-30)

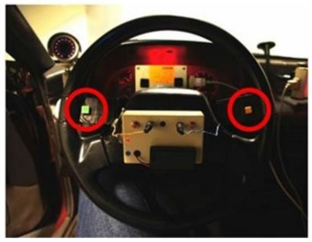


Figure 2-30. Steering Wheel

The study design, shown in Figure 2-31, features three conditions with different stimulus onset to investigate neural correlates of shifts of attention between the lane-departure events and a dynamic attention shifting event. The lane-keeping driving task in Case 1 was single-task. Case 2 and Case 3 were dual-task; two tasks simultaneously onset. The deviation onset and the visual target stimulus appear in Case 2. The deviation onset and the auditory target stimulus in Case 3. On average, there were 60 occurrences of each case during every 15-min experimental session.

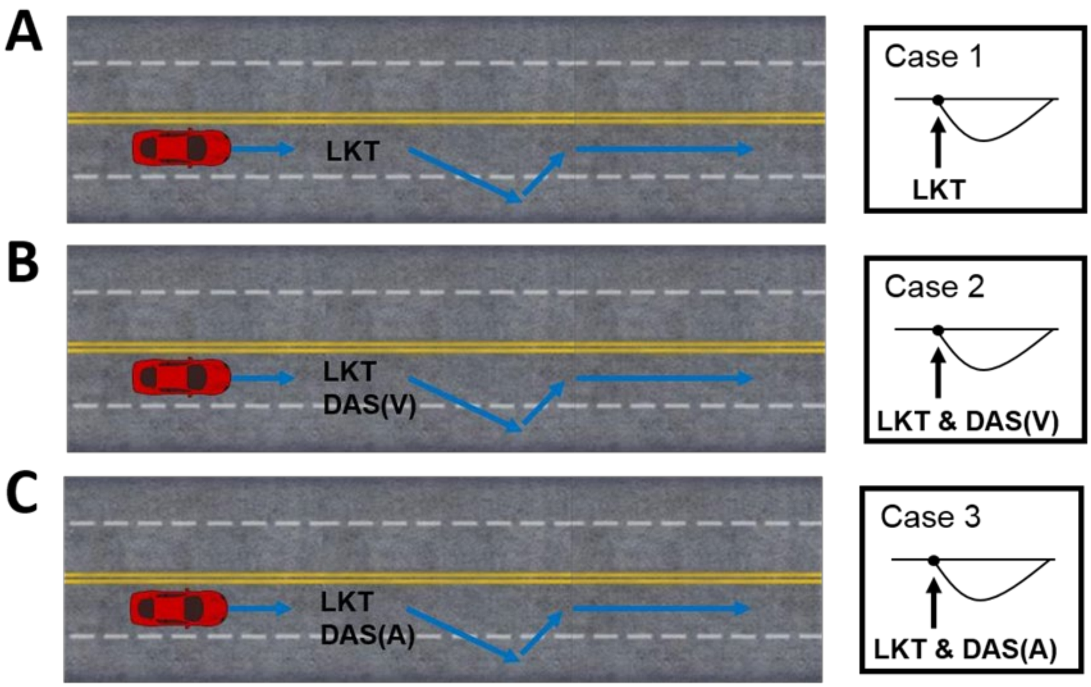


Figure 2-31. RWN-VDE Experimental Conditions

2.4.1.4.7 POC

Chin-Teng Lin (ctlin@mail.nctu.edu.txw), NCTU, Principle Investigator
Scott Kerick (scott.e.kerick.civ@army.mil), ARL, Research Collaborator
Dave Hairston (william.d.hairston4.civ@army.mil), ARL, Research Collaborator

2.4.1.5 TNO FLERP Experiment (TNO_CANCTA_FLERP)

The p300 event related potential (ERP), a positive peak occurring in the EEG signal roughly 300 ms after a sensory event, indicates that an observer's attention has been drawn. It has been shown to reliably distinguish between top-down defined 'targets' and 'non-targets', even on a single ERP basis, e.g. in cases where observers are asked to pay attention to the letter 'p' presented in a stream of successively presented letters. However, in typical P300 studies and (Brain-Computer Interface) applications, target and non-targets are imposed to the observers who are asked to not move their eyes. In contrast, observers actively and purposefully move their eyes in most real world search or monitoring tasks. We are interested in using fixation- rather than stimulus locked ERPs (FRPs) as a means to determine whether observers are looking at a target (i.e. a relevant object or not). Few studies have examined fixation-locked late ERPs, but both ARL and TNO already performed several studies in this respect. In Brouwer et al. (2013, 2014) it has been shown that distinguishing between target and non-target fixation is possible above chance on a single FRP basis, even when controlling for potentially confounding factors such as saccade length and low-level visual features.

2.4.1.5.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL/TNO
Protocol Number: ARL 15-128
Protocol Name: "Fixation-locked EEG"
Contract: W911NF-10-D-0002-0020

2.4.1.5.2 Location

The study was conducted at the TNO laboratory (Netherlands).

2.4.1.5.3 Subjects

Twenty-one participants (nine males, twelve females) between the age of 19 and 30 (average age: 22, 9) were recruited through the participant pool of the Netherlands Organization for Applied Scientific Research (TNO). None of the participants wore glasses. Each participant received a monetary reward for his or her time and travel costs. All participants signed an informed consent form preceding the experiment. This study was conducted in accordance with the Army Research Laboratory's IRB requirements (32 CFR 219 and DoDI 3216.02).

2.4.1.5.4 Apparatus

The task was presented on a 19-inch flat-screen monitor (Dell 1907FP Flat panel 19"). The screen resolution was 1280x1024 and the refresh rate was set at 60 Hz. Participants were located approximately 40 cm from the screen. Audio output was coming from a dual speaker set (TEAC PowerMax 60/2) placed left and right of the screen.

Gaze and pupil size were recorded at 60 frames per second using SmartEyePro V6.1.6 (Smart Eye AB, Göteburg, Sweden). This system consists of two cameras (Basler acA640-120gm, HR 8.0 mm lens) placed at the left and right side of the screen.

EEG and EOG signals were recorded using an ActiveTwoMK II system (BioSemi, Amsterdam, Netherlands) with a sampling frequency of 512 Hz. For EEG, 32 active silver-chloride EEG electrodes were placed according to the 10-20 system and were referenced to the Common Mode Sense (CMS) active electrode and Driven Right Leg (DRL) passive electrode. Four EOG electrodes (BioSemi Flat Active electrodes, Amsterdam, Netherlands) were used to record eye movement. Two EOG electrodes were placed at the approximately 0.5 cm off the lateral canthi of both eyes, and were used to record horizontal eye movement. Another two EOG electrodes were placed above and below the left eye to record vertical eye movement and blinks. The impedance of all electrodes were <25 kΩ.

2.4.1.5.5 Demographics

Demographic information provided with this dataset includes subject ID, gender, and age.

2.4.1.5.6 FLERP Tasks

The experiment features two tasks: a monitoring task and an auditory math task. In the high load condition, participants performed both tasks. In the low load condition, they only needed to perform the monitoring task, even though the math task was still played to keep auditory stimulation constant across conditions.

<u>Monitoring task</u>:

Participants were asked to monitor 15 systems, represented by strings of symbols on a screen and placed in three rows of five columns. There were three different system conditions: hidden ('####'), working as intended ('#OK#') or system failure ('#FA#'). At the start of a trial, all system conditions were hidden. Then, each of the systems was successively highlighted for 1s (1027 ms) by displaying a square around it while its condition changed from '####' into either '#OK#' or

'#FA#' (see Figure 2-32). Highlighting the systems happened in random order, except for that two subsequently presented systems were never further apart than two steps in horizontal direction and one in vertical direction, or two vertical and one horizontal. The next highlighted system was always in peripheral vision such that it was impossible to distinguish between '#OK#' or '#FA#' without making a saccade. After all system conditions had been shown, empty boxes appeared at the system locations and the participant had to indicate which systems failed during the trial by clicking the appropriate boxes with the left mouse button. When finished, the participant pressed the OK button at the top left of the screen. Every trial, two, three or four '#FA#'s were presented. The amount and the '#FA#' locations were chosen randomly.

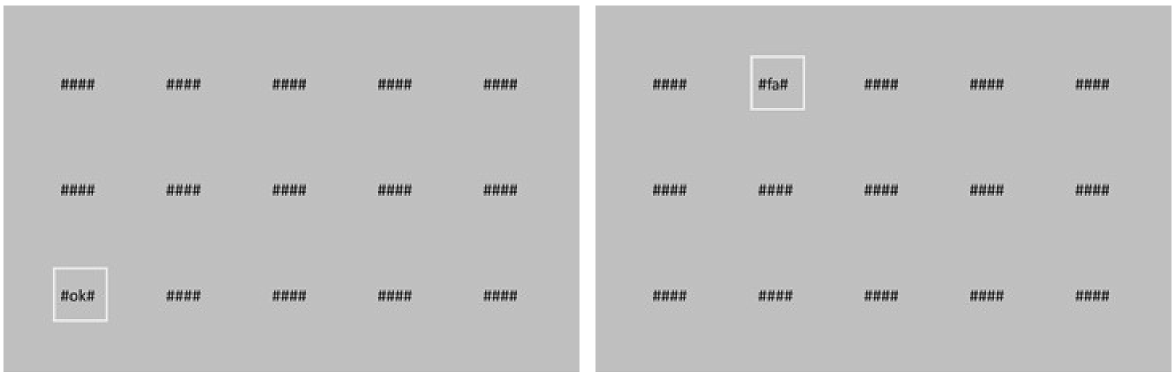


Figure 2-32. FLERP Monitoring Task

Math task:

The math task was an aurally presented sum consisting of six numbers between 6 and 12. Only addition (+) and subtraction (-) operations were used. The first number was presented one second after the start of the monitor task, and every 2660 ms another number was presented. Thus, the last number was presented after 14.3 seconds. When participants had to perform the math task (i.e. in the high load condition), they were required to give the answer of the sum after having indicated where the '#FA#'s were located. This was done by typing the answer and pressing enter. In order to motivate participants to perform the math task so that the conditions would actually differ, they received feedback on their answer. If the answer was incorrect, the correct answer was shown.

For each of the load conditions, participants performed 8 blocks of 11 trials. High and low load conditions were presented alternately, starting with the high load condition.

2.4.1.5.7 POC

Anne-Marie Brouwer (anne-marie.brouwer@tno.nl), TNO, Primary Investigator
Jon Touryan (jonathan.o.touryan.civ@army.mil), ARL, Associate Investigator
Anthony Reis (anthony.j.ries2.civ@army.mil), ARL, Associate Investigator
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer

2.4.1.6 TNO ACC Experiment (TNO_CANCTA_ACC)

Currently, vehicles behave the same way irrespective of most environmental circumstances (e.g. traffic density) or the driver's mental state. For instance, Adaptive Cruise Control (ACC) systems (adaptive in the sense that they respond to vehicles in front) decelerate following a fixed velocity profile when you are approaching a vehicle driving at a lower speed. However, when traffic is busy, you are stressed and you do not want other cars to sneak in between you and the vehicle in front, you will probably prefer a stronger deceleration than when you are relaxed, on an empty highway, and your mother is in the passenger seat. Our ultimate aim is to integrate the driver into the system such that comfort is enhanced within safety margins. This requires the use of information about the driving environment, the vehicle and the driver to adapt ACC settings in real time. ACC makes a good case for responding to the estimated wishes and intentions of the driver since 100% accuracy would not be required to create a system that is closer to what drivers like than one that is not adaptive in this sense at all. To make this possible we combine our expertise in developing and testing ACC systems (TNO Helmond), monitoring and affecting driving behavior using vehicle parameters and environmental variables (TNO Soesterberg, ARL, Zander Laboratories) and estimating cognitive and affective state through EEG and other physiological variables (TNO Soesterberg, ARL, Zander Laboratories).

2.4.1.6.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL/TNO
Protocol Number: ARL 16-038
Protocol Name: "Detecting Unexpected Adaptive Cruise Control (ACC) Behavior"
Contract: W911NF-10-D-0002-0026

2.4.1.6.2 Location

Data collection for this experiment took place at the RDW Testcentre in Lelystad, the Netherlands (Figure 2-33). The facility featured a circular dedicated track (100m radius).

https://www.rdw.nl/sites/tgk/englishversion/Paginas/Test-Centre-Lelystad-(TCL).aspx?Path=Portal/TGK/English%20version/Testing

Figure 2-33. ACC Test Course

No other traffic was present during the ACC task. An experimental leader was present in the backseat of the car during the whole experiment.

2.4.1.6.3 Subjects

Fifteen participants (9 men, 6 women; age range: 24-60 years) were recruited from the TNO research participant database to take part in the experiment. Participants possessed a driving license for at least three years. They received a monetary reward to make up for their travel and time. The study is in accordance with the Declaration of Helsinki and has been approved by the local ethics committee. All participant signed an informed consent form prior to taking part in the experiment.

2.4.1.6.4 Apparatus

The experimental vehicle (shown in Figure 2-34) was an instrumented Toyota Prius (TNO Helmond), capturing vehicle state and driver performance data via the CAN bus, and MobilEye system.

Figure 2-34. ACC Instrumented Vehicle

Subjects were instrumented with a BioSemi 64 (+8) channel EEG system, as depicted in Figure 2-35. EEG electrodes were placed according to the International 10-20 system. For better signal understanding and to facilitate artefact removal, the experiment team also recorded the following modalities, collected through the external channels (non-scalp) of the BioSemi data acquisition system:

1. EOG (electrodes were placed above and below the left eye of the participant)
2. EMG (electrodes where placed at the neck of the participant, 2 on the left trapezius muscle and 2 on the right trapezius muscle, 3 cm above each other)
3. ECG (sensors at right collarbone and the lower left rib).

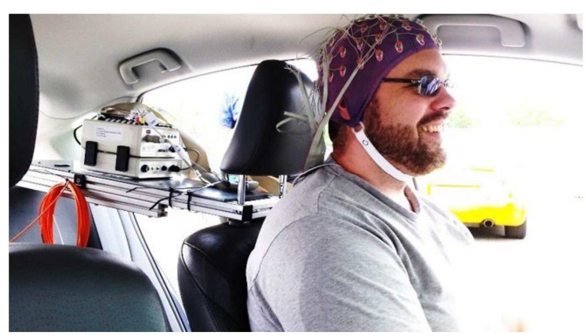


Figure 2-35. ACC Engineering Test Participant

All physiological signals were recorded at a sampling rate of 512 using a BioSemi amplifier (USA, BioSemi Active-Two)

2.4.1.6.5 Demographics

Demographic information provided with this dataset includes subject ID, gender, and age.

2.4.1.6.6 ACC Tasks

Participants were tested individually in the car and asked not to make unnecessary movements. They started with 10 practice trials to get familiar with the task and settings, followed by 300 experimental trials. Participants were told that we are working on detecting a driver's desired deceleration settings of an ACC, without requiring the driver to communicate this desire explicitly. We told them that in this stage, we cannot do that yet and therefore, we represent the driver's wish concerning either strong or soft braking by presenting a human voice stating the desired setting.

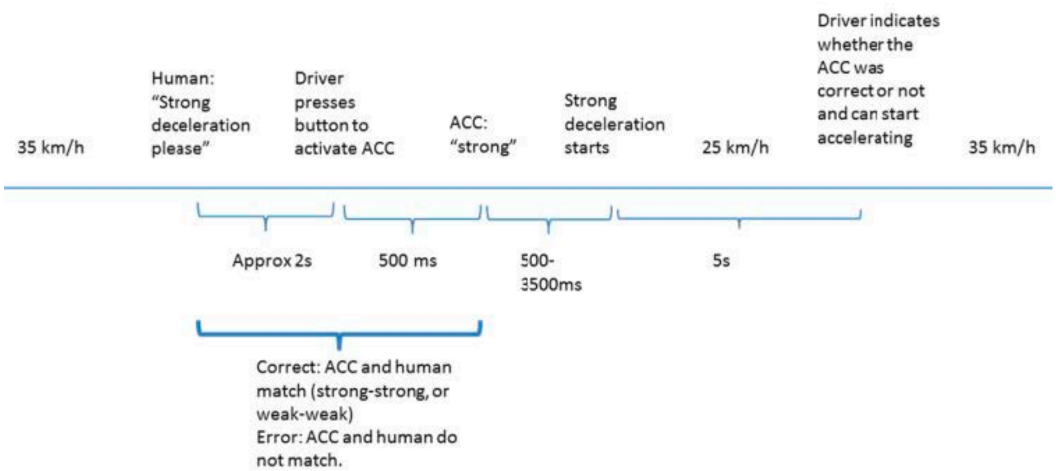


Figure 2-36. Events Comprising an ACC Experimental Trial

The car drove at 35 km/h on the track when a human voice indicated the desired deceleration (strong or weak). The participant pressed a lever down to activate the ACC deceleration. Before the deceleration, the ACC announced through a computer voice whether it would decelerate strongly or weakly. In 80% of the trials the wish of the human voice was followed (match trial), but in 20% the deceleration profile did not match the human voice (mismatch trial). A variable time between 0.5 and 3.5 s after the ACC's announcement the car decelerated to 25 km/h, following a steep or a shallow velocity profile (strong or weak deceleration, with a maximum deceleration of respectively 3 or 0.7 m/s^2; where going from 35 to 25 km/h took respectively 0.9 or 2.8s). Then the human voice asked the driver to accelerate again. The driver indicated whether the ACC had followed the desired type of deceleration or not, and pushed the lever up to have the ACC accelerate to 35 km/h. This series of activities is represented as a timeline in Figure 2-36.

The vehicle braked strongly in half of the trials, and softly in the other half. After every 40 experimental trials there was a short break to relax and turn around the car into the other direction. The total drive lasted for about one and a half hours.

2.4.1.6.7 POC

Anne-Marie Brouwer (anne-marie.brouwer@tno.nl), TNO, Primary Investigator
Kaleb McDowell, ARL, Associate Investigator
Oded Flascher, DCS, Associate Investigator
Matthew Jaswa (mjaswa@dcscorp.com), DCS Corp., Systems Engineer
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer

2.4.1.7 Assessment of Intra- and Interpersonal Brain and Behavioral Dynamics of Driver-Passenger Dyads during Real-World Driving

This project is referred to using the identifier RWN-VDEDP (Real-World Neuroimaging in Vehicle Driving Environments with Driver and Passenger).

In military and civilian domains, human drivers are gradually being replaced by automated systems. This is precipitating a shift towards humans experiencing vehicles from the primary perspective of being a passenger. While decades of research have provided vast amounts of data regarding vehicle drivers, very little is known about how people experience being a passenger in a vehicle that they do not control. Whether driven by another human or an automation, passengers have regular opportunities to make decisions regarding ongoing driving performance, and in many cases these decisions lead to interventions. Such passenger behaviors may critically impact vehicle behavior and, more importantly, when and how they occur is likely to be subject to a variety of factors such as individual decision-making style, driving preferences, level of fatigue, risk tolerance, and personality (French et al, 1993; Recarte & Nunes, 2003; West, Elander, & French, 1992). In particular, human decisions regarding cooperative behaviors such as reliance and compliance, are believed to be significantly impacted by the amount of trust held in the paired agent, whether human or automated (Madhavan, Wiegmann, & Lacson, 2006; Parasuraman & Riley, 1997), and some individuals are more susceptible to peer influence when prompted to adopt risk-taking behavior (Wasylyshyn et al., under review). Therefore, it is important to understand what individual state and trait factors lead to passenger trust in a vehicle operator as these are likely to inform the design of intelligent communication systems supporting future human-agent teaming solutions that will address long-term Army Warfighting challenges.

Among the many factors believed to be important to a human's ability to trust is the degree to which the agent to be trusted (trustee) has intentions that are transparent and communicated in a manner that is understandable (c.f. Barnes, Chen, & Hill, 2017). This indicates a cognitive component to trust that develops in the context of an emotional tone, or valence, defined within the trustor-trustee interaction (Borum, 2010). A prominent line of inquiry that may be brought to bear on the formation of trusting relationships is how emotional valence, such as that induced by a relatively positive versus neutral tone, can enhance the transfer and retention of information through communication (for a review, see Delli Carpini, 2012). We adopted a working hypothesis that positive emotional valence in human interactions would likewise improve communication efficacy, as evidenced by strengthened learning and memory, and this increase in positive communication might also lead to increased perception of trust. More specifically, we operationalized positive emotional valence through the experimental manipulation of "humor" and then hypothesized that psychological processes involved with intrinsic reward (e.g., satisfaction of hearing a joke) as well as the desire to share the information with others, would precipitate positive effects that enable more effective communication. This hypothesis was based on previous studies showing that humor modulates reward centers in the brain and that reward-related learning enhances memory (Adcock, et al., 2006; Mobbs, et al., 2003). Though humans and driving automations are not direct analogues, it has been shown that people have a strong tendency to attribute human motives to non-human objects (Heider & Simmel, 1944). Moreover, emotional context of interpersonal interactions can have an impact on communication effectiveness, but research has shown that an individual's sensitivity to emotion can be strongly influenced by recent sleep loss (Vandekerckhove and Cluydts, 2010; Wong et al., 2013). However, the interaction of sleep loss and emotional state fluctuations is not well understood, particularly in naturalistic contexts (Walker and Harvey, 2010). Thus, monitoring long-term sleep patterns and daily estimates of the emotional valence of social interactions provided a novel approach to understand how such factors influence passenger perceptions of driver trustworthiness. By better understanding the influence of naturalistic fluctuations in state across days and weeks and its interaction with trust development, our research enables future extension to examining how such findings may translate to improved passenger trust in driving autonomy.

We developed a novel paradigm that allowed us to track driver-passenger communication as a function of emotional valence, here manipulated by use of humor, and whether or how these factors influenced the development of passenger trust in the driver. Further, we stepped deeper into these socio-psychological processes by examining the individual and joint brain and peripheral physiological activations that reflect the changing individual and interpersonal interaction dynamics. Prior research demonstrates that brain activity during basic social tasks can predict driving risk and susceptibility to social influence (Cascio et al, 2015; Falk et al, 2014; Schmälzle et al, in press; Wasylyshyn et al, under review). Therefore we experimentally manipulated emotional valence (humorous vs. neutral framing) within subjects and measured the degree to which that valence alters behavioral indicators of passenger trust in the driver's capability (Aim 2, below), success of communicating with the driver (Aim 3), and how these relationships are altered depending on the driver-passenger interaction frequency outside of the experimental protocol as well as broader characteristics of each participant's general social network (Aim 4). Finally, across all of these aims, we sought to understand how performance is influenced by naturalistic changes in an individual's cognitive state, indexed by sleep history, physical activity, and daily estimates of social interactions and risk tolerance. This project leveraged a unique combination of an instrumented vehicle, the manipulation and measurement of interpersonal communication dynamics, and rich multi-modal streams of physiological data captured in a real world driving task with high ecological validity. This interdisciplinary approach thereby provided new insights into the social and psychological experiences of vehicle passengers while traversing actual roads in natural traffic conditions, translating our research from simulated laboratory environments to more holistically-considered operational settings reflecting challenges to be addressed in top priority Army technology areas such as man-unmanned teaming in complex urban environments.

This research aimed to gain new insights into problems associated with real-world, role-based human variability during driver-passenger interactions in a vehicle traveling on typical roadways in regular traffic. Both inter- and intra-individual variability were of interest; particularly factors associated with state variables, such as the communication tone, as well as more persistent trait influences including individual biases or proclivities (e.g., propensity to trust) and those from the individual's social network. Further, we were interested in inter- and intra-individual variability that arises from naturalistic fluctuations in cognitive state, indexed by sleep history, physical activity, daily estimates of quality of social interaction, and daily estimates of risk preferences. The aims for this goal included:

Aim #1. We investigated whether our previous inferences regarding brain-behavior relationships from simulated driving environments translate to natural task environments.

Hypothesis 1a: Using a variant of Granger causality (Garcia et al., 2017), we expected that driving behavior would predict brain activity during time periods of quiescence in the drive, whereas brain will predict driving performance when the driver is navigating traffic, making route decisions, or listening to the passenger. More specifically, we expected that news stories that the driver remembers well (indexed by correct memory on the post-drive task) will correspond with time frames when the brain is less tightly coupled with the driving performance and/or different brain regions are engaged than the ones identified from the simulated environment where the driver was alone.

Hypothesis 1b: We expected that drivers who had recent sleep loss and/or low levels of physical activity would have weaker coupling between the brain activity and their driving and listening tasks. That is, the driver will have increased timeframes where the driving behavior predicts their brain activity based on decreased abilities to focus on non-primary tasks.

Aim #2. We assessed the effect of emotional valence in conversation, as operationalized through humor, on trust in a driver.

Hypothesis 2a: Exposure to and engagement with humorously framed material would increase the level of trust between passenger and driver compared to material with neutral framing. More specifically, through both covert behaviors and overt ratings, passengers who are more responsive to the humor manipulation would likewise show higher trust in the driver's performance as compared with passengers that respond equally to humorous and neutral material.

Hypothesis 2b: We expected that drivers who reported daily social interactions that have been positive and rewarding would have an accurate memory for the news stories told by the passenger and increased ratings of trust.

<u>Aim #3</u>. We examined the degree to which emotional valence impacts success of communication.

Hypothesis 3a: Given that many key tasks require coordination between team members, one of whom may be operating a vehicle, it is critical to understand how successful communication is accomplished in this environment. We predicted that information framed humorously would be more successfully communicated between passenger and driver than information framed with neutral valence. To demonstrate this, we examined three sub-hypotheses:

a) Memory for information in the humorous framing condition would be greater than the neutral condition.
b) Increased synchrony in neural responses to information would be observed between driver and passenger during the drive when originally presented with humorous (vs. neutral framing).
c) We predicted that greater levels of inter-subject correlation (ISC; Hasson et al, 2008), as measured using whole brain EEG, would be associated with higher accuracy on the post-drive memory tasks.

Hypothesis 3b: We expected that drivers who had reported daily social interactions that have been positive and rewarding would have increased ISC with their passengers. Likewise, we expected that drivers with minimal sleep loss and higher activity would remember more stories and score higher on the memory tasks.

<u>Aim #4</u>. We examined how individual differences in social network structure are associated with the success of communication during the drive.

Hypothesis 4a: Based on our recent work (Schmälzle et al, 2017), we expected that participants with more dense networks would show increased memory performance, as they have a social structure that reflects tight connections where information is routinely brokered among colleagues and therefore requires successful listening and retention skills. As a more specific analysis of the dyad, we also indexed the frequency of interaction between the driver and passenger, and again, we expected that previous connection would increase the success of communication, indexed by trust ratings, higher inter-subject correlation during in-drive communication tasks, and memory performance on the post-drive tasks.

Hypothesis 4b: We expected that individuals with more dense social networks would be more willing to take risks, and this would be reflected in the daily estimates of risk tolerance. Concomitantly, given the connection between trust and risk, we anticipated individuals who are more risk tolerant to show greater propensity to trust as well as higher ratings of the trustworthiness of their dyad partner.

2.4.1.7.1 Protocol

The research protocol information for this study is as follows:

Organization: ARL
Protocol Number: ARL 17-092
Protocol Name: "Assessment of Intra- and Interpersonal Brain and Behavioral Dynamics of Driver-Passenger Dyads during Real-World Driving"
Contract: W911NF-10-2-0022

2.4.1.7.2 Location

Data for this research was collected at US Army Research Laboratory-Human Research and Engineering Directorate (ARL-HRED) using a road network in the greater Harford County area that is patrolled by Maryland State Police, Barracks M.

2.4.1.7.3 Subjects

A total of 44 subjects, paired into 22 dyads, participated in this study. In all but one dyad both subjects acted as driver and passenger during different sessions on different days. Only one of the subjects acted as driver for the 44th dyad, however for two other dyads, the EEG data were corrupt. Therefore a total of 40 trials sessions were conducted that yielded useable EEG recordings and useable vehicle data. 92 total useable EEG recordings were produced, 42 from the driver subject and 50 from the passenger subject.

2.4.1.7.4 Apparatus

<u>*Testbed Vehicle:*</u> The primary location of data collection was within a standard commercially-available vehicle that was leased for use in this research (Figure 2-37). The specific vehicle for this experiment was a 2016 Ford Fusion Titanium (http://www.ford.com/services/assets/Brochure?bodystyle=Sedan&make=Ford&model=Fusion&year=20164). The leased model included the standard Ford 2.0L EcoBoost® I-4 engine, automatic transmission, 4-wheel anti-lock braking system, electric power-assisted steering, and a full suite of safety features.

Figure 2-37. Stock Interior and Exterior Views of the 2016 Ford Fusion Titanium

ARL MARIN System: The ARL Multi-Aspect Real-world Integrated Neuroimaging (MARIN) system was designed to leverage and integrate COTS technology to enable monitoring of psycho-physiological, behavioral, and subjective experience changes in real-world, daily life environments. The MARIN was validated, with user assessment, through technology evaluation that was approved by administrative review of the ARL IRB Chair as described in protocol # ARL 13-036 ("The MARIN System: Initial Usability and Technology Validation", P.I. T.J. Doty). The MARIN was configured for this experiment to enable logging of electroencephalographic (EEG), electrocardiographic (ECG), electrodermal (EDA), photoplethysmography (PPG), respiratory, and physical activity, and skin temperature. The components of the system included:

- EEG: *ABM B-Alert X24 System*

The B-Alert X24 Wireless EEG system (Advanced Brain Monitoring, Inc, Carlsbad, CA) uses a flexible electrode strip that was affixed to an adjustable headband and fit snugly to the head of the participant. The flexible strip houses a set of flat electrodes in standard international 10-20 scalp locations. Cylindrical foam pads were moistened with a conductive paste then placed under each electrode to serve as the conductive medium between scalp and sensor. In addition, as part of the B-Alert headset, a wireless transmitter was attached to the headband at the back of the head allowing complete freedom of head movement as compared with a standard wire-tethered system. This transmitter sent the EEG data to a separate device for data logging. The weight of this system was less than 200g.

- ECG, respiration, movement, and posture: *Zephyr Bioharness 3*

The Bioharness 3 (Zephyr Performance Systems, Annapolis, MD) is a small, lightweight (50 g) device that was worn around the upper torso, directly against the skin of the participant. The Bioharness enables the capture and wireless transmission of peripheral psycho-physiological data of the wearer, including medical grade ECG, respiration, movement, and posture, without requiring tethering to external data collection equipment or any other hindrance to free movement.

- In-Drive EDA, PPG, skin temperature, and actigraphy: Empatica E4

The Empatica E4 (Empatica, Inc., Boston, MA) is a small, lightweight, wrist-borne device for real-time capture and wireless transmission of psycho-physiological data. The E4 contains four embedded sensors that provide data regarding PPG, EDA, skin temperature, and physical activity (actigraphy) through 3-axis accelerometry. The device has the ability to operate in data streaming mode as well as in an in-memory mode enabled by internal flash memory for continuous data recording when away from a data logging device.

- Daily Actigraphy for Sleep and Physical Activity: Readiband

The Readiband is a lightweight, wrist-borne device that measures sleep and physical movement through high-frequency 3-axis accelerometry. The device uses the proprietary SAFTE Fatigue Model to classify sleep and wake states. The device may be worn for 30 days on a single charge and stored all logged data until downloaded for storage at ARL.

- Mobile Data Recording

Data logging in the MARIN system was enabled by a custom programmed Android-based mobile device that could be easily carried with the participant or mounted in the vehicle. The mobile device received and stored data streamed from the psycho-physiological devices.

A B

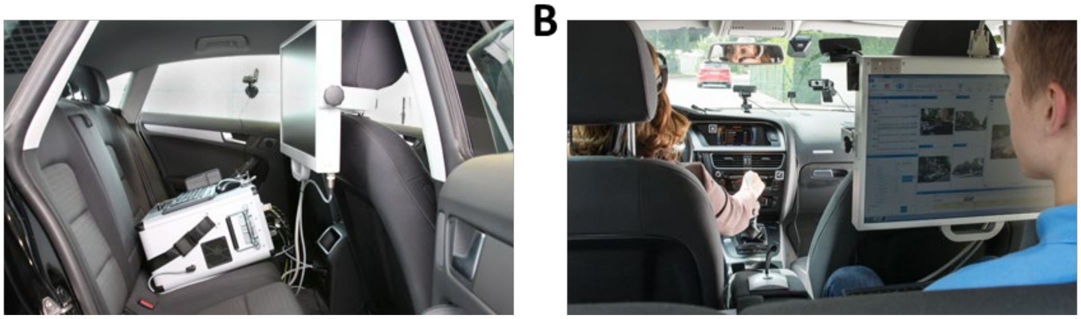


Figure 2-38. The Automotive Computer Shown Securely Mounted on the Rear Bench Seat (A), System Monitor Shown Securely Mounted to the Rear Headrest of the Passenger Seat (A, B)

Ergoneers Vehicle Test Kit: Figure 2-38 shows the components of the vehicle test kit (Ergoneers GmbH, Manching, Germany), which is a COTS system for conducting natural driving studies. A single Automotive Computer comprises a complete system for data recording and analysis. The vehicle test kit allows integrated and synchronized acquisition of data from humans (i.e. video and audio), vehicle behavior (i.e. vehicle Controller Area Network, or CAN, interface, a Mobileye® camera-based lane-position tracking, GPS), and custom inputs. System components include:

- Video/Audio Recording System

Video data was captured using 4 cameras mounted on the inside and around the testbed vehicle in positions that did not obstruct the driver's view of the roadway. Two cameras faced into the cabin, one trained on each participant, and two captured external environment images. All video data was primarily used to facilitate interpretation of later quantitative data. Audio data was recorded in order to capture dialogue between the participants and the audio cues. Participants were made aware of the recording and archival of the video and audio data, as well as steps taken to prevent the release of any personally identifiable information.

- Passenger Seat Sensors

Passenger postural responses during the drive were captured through an array of force-sensing resistors embedded in a fabric cover fitted over the passenger seat. The individual inputs from the force-sensor array were multiplexed and integrated through a custom Arduino card, which then transmitted continuous voltage outputs to the test kit Automotive Computer through a single USB connection.

Figure 2-39(left) provides a depiction of the configuration of the two arrays of force sensing resistors, and numbering corresponds with labels inside the data files for `Seat` and `SeatBack` variables.

Figure 2-39 Configuration of sensor arrays on the passenger seat

- Standard PCs, Monitors, and Input Devices

In addition to the Ergoneers Automotive Computer, this project required the use of a Microsoft® Surface Book computer for passenger task and stimulus control and two Smartphones running the Android operating system (see MARIN, above). Each device recorded a common vehicle on-board diagnostic (OBD) CAN bus signal to enable synchronization across all data streams (MARIN inputs included). The monitor of the vehicle test kit was mounted securely to the headrest of the passenger seat, and made use of a wireless keyboard for input. Tablet Surface Book and Android devices made use of their native input methods (built-in keyboards and touch screens).

Figure 2-39 provides an integrated, systems-level overview of the various apparatus and equipment that was used to support this experiment.

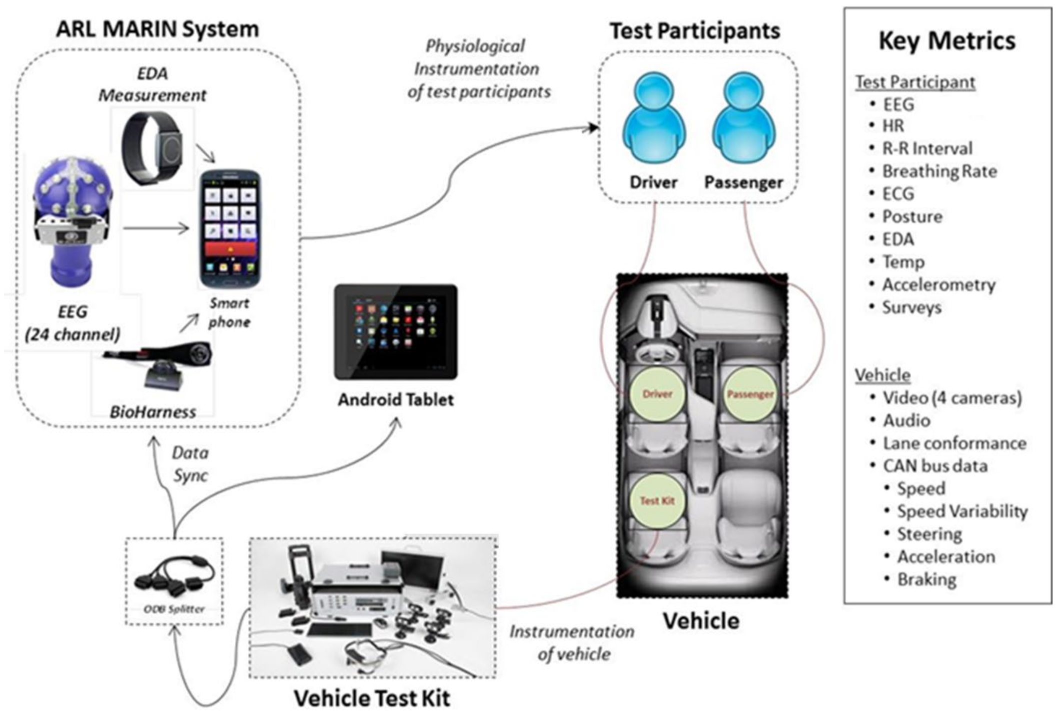


Figure 2-40. Overview of System Developed to Support the Proposed Experiment

2.4.1.7.5 Demographics

Demographic information collected for this study includes subject ID and age, however, these data were not processed in the current version of the data set. For this reason, this demographic data not currently available in the public raw data either; it is available, but would need to be mined from the Participant Information Questionnaire.

2.4.1.7.6 RWNVDEDP Tasks

Figure 2-40 provides a timeline of the participant experience in this experiment. The top blocks in the figure indicate weeks wherein the participants had daily text-message based experiment interactions interspersed with the driving sessions; the lightly shaded squares indicate potential variability due to scheduling. The blue shading indicates when the short social interaction questions were asked throughout and the yellow and orange shading in the Interim and Follow-up periods indicate two different additional reminders depending on the content of the JAVI task from the corresponding drive (physical activity or sleep). The final square along the top indicates that Close-out happened virtually by secure URL. The timeline for the two driving sessions is indicated underneath, as indicated by the blue arrows. Session time begins at the top and progresses downward. PA = Passenger timeline, DR = Driver timeline.

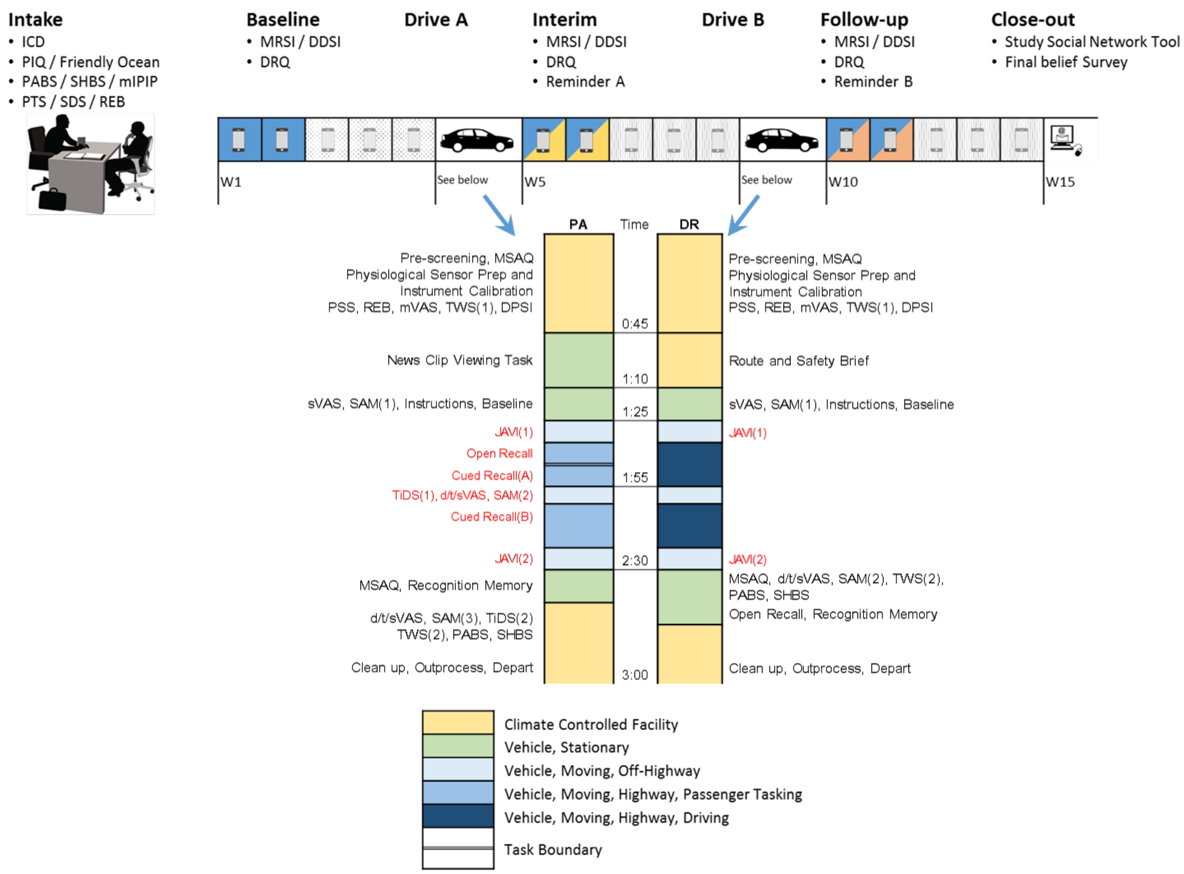


Figure 2-41. RWNVDEDP Timeline

Intake

When the individual volunteered to be a participant in the study, they were first asked to complete a separate 30-to-45-minute intake session. This intake session began with the informed consent dialogue between experimenter and participant (guided by the Informed Consent Document, ICD). After providing fully informed consent, the participants completed a set of questionnaires and psychometric tests. First, the Participant Information Questionnaire (PIQ) was administered to gather a limited set of information about demographic factors that could influence the results. Questions included standard demographic information (age, occupation, education) and driving experience. In addition to individual demographics, we were also interested in broader social influences and beliefs regarding health behaviors.

A tool developed in the Falk Lab at UPenn was used to gain a better understanding of each individual participant's relationship with their broader social network. This tool, known as Friendly Ocean, included questions about their social interactions online, in person, and over telephone and text messaging. We also asked specific, targeted questions to enable characterization of the general relationship between the dyad partners (Driver-Passenger Social Interaction Questionnaire, DPSI) and, at the end of all data collection, among all study participants who opted-in to inclusion in a Study Social Network Information Tool. Participants also provided responses to two brief questionnaires regarding their general daily social interactions. These included 2-3 questions each about their Most Recent Social Interaction (MRSI) and their Daily Diary Social Interactions (DDSI). To gather information regarding the beliefs participants held regarding health activities (physical activity and sleep habits), they were asked to respond to PABS and SHBS surveys based on the work of Fishbein and Raven (1962). These surveys were repeated at three additional points in the experiment (after each drive and at close-out).

Beyond demographic factors, we also used instruments that provided measures of general aspects of each individuals' personality. The mini-IPIP (mIPIP; Donnellan et al., 2006) provided insight into the larger scope of the human-related factors that influence trust and communication. In particular, the mIPIP efficiently indexes the well-established "Big Five" factors of personality, including extraversion, openness to experience, agreeableness, conscientiousness, and neuroticism. To obtain more direct measures of personality traits related to interpersonal trust, we also administered a Propensity to Trust (PTS) and Trustworthiness scales as described in Evans and Revelle (2008), which were drawn and validated using items from the International Personality Item Pool (Goldberg et al, 2006). The particular comparisons of interest in this case were the PTS for each participant as well as the stability of their perception of the trustworthiness (TWS) of their dyad partner. To provide a basis of comparison for the TWS, we also had participants respond to questions from the short form of the Social Desirability Scale (SDS), which indexes whether the participants are likely to be responding in 'culturally sanctioned ways'. Finally, at an initial intake session and at each of the driving sessions, the participants completed a set of questionnaires to elicit risk preferences, including three about non-social risks and three about social risks using the Risk Elicitation Battery (REB). The sub-items of the REB also provided the content for a daily indicator of risk orientation when not in the driving sessions. That is, in addition to the MRSI and DDSI, participants answered one question each day regarding their risk preference (the Daily Risk Question, DRQ, refer to protocol document for details).

The investigator oriented the participants to the components of the experiment involved in daily monitoring and then obtained the participants' preferred phone numbers, usual wake up times, and usual sleep time to allow for message scheduling. Next, the participant received an actigraph watch (Readiband) and was given an explanation about what it measures.

Baseline Period

There was a 2 – 5 week baseline period between Intake and Drive A. During that time, the participants wore the Readiband, and the data collected was used to establish participants' baseline levels of physical activity and sleep habits. Participants received two text messages each day that included a link to an online set of risk and social interaction questions accessible through a smartphone browser. The first daily text, which was sent sometime in the middle of the day based on the participant's self-reported preferred wake up time, linked the participant to the MRSI and the second text, sent each evening, linked the participants to the DDSI. There was also a single risk-elicitation question (the DRQ) with each prompt. The data collected during these weeks was used to establish participants' baseline social interaction tendencies.

Drive A

Prior to sensor preparation, each participant answered the questions on a custom-designed participation-readiness questionnaire to assess their general fitness to participate in roadway driving on that day. Once the readiness was confirmed, then participants provided baseline responses to the MSAQ. Participants were then outfitted with their set of physiological sensors in a climate-controlled space and each underwent a role-specific set of preparatory steps (see the vertical columns in Figure 2-40; PA = passenger and DR = driver). While the experimenters were placing the sensors on the participant and adjusting their calibration the passengers completed several single administration subjective instruments including the DPSI, REB, Perceived Stress Scale (PSS), and motivation Visual Analog Scale (mVAS). Both participants also completed a pre-drive TWS, a baseline stress scale (sVAS), and the Self-Assessment Manikin (SAM).

The PSS is a short, 10-question inventory designed to assess the current level of stress of a given individual. The PSS indexes the general level of stress that an individual has been feeling for approximately the prior month. This inventory was used to provide data that would account for individual differences in the baseline level of stress that an individual participant had at the outset of this study.

Visual Analogue Scales (VASs) efficiently provide indices of subjective state changes in response to the experimental manipulations. VASs are sliding scales anchored by the appropriate text descriptors at the two extremes of the slider. In general, VASs provide a visual guide for participants to rate their subjective perception of specific feelings or characteristics on a continuous scale by choosing a location between two extremes. For the current study we used such scales to index stress (sVAS), general motivation (mVAS), drive difficulty (dVAS), and traffic density (tVAS).

The SAM was developed to provide an efficient and reliable method for assessing emotional state along the dimensions of pleasure, arousal, and dominance (Bradley & Lang, 1994). Here, we used a modified version that provides additional "anchoring" terms to assist the participants in selecting their responses (based on Schifferstein, Talke, & Oudshoorn, 2011). Similar to the well-known PANAS, this tool was intended to provide evidence of qualitative differences in basic dimensions of emotional state throughout the experiment.

After the setup was complete, the passenger was escorted to the vehicle where, using a tablet PC, they watched and rated the likelihood they would share each in a set of 16 news stories (half humorous, half neutral, randomly sequenced). Concurrently, the driver reviewed the route, vehicle controls, and safety procedures with an experimenter. Upon completion of their respective tasks, the driver and passenger both sat in the car and watched a short video that reviewed the flow of the tasks during the drive, the set of post-drive tasks, and discussed the role of the staff member. Next, the participants both sat comfortably with their eyes open in a resting state condition to record baseline measurements of their physiology. The participants were instructed to remain silent and looking forward during this period.

A basic safety check was performed and the drive began, initiating the Open Course Driving task for the driver:

- Task: During each drive, the driver listened to the passenger describing previously watched videos or a pre-recorded and amplitude modulated podcast (JAVI, described below). The driver was asked to hold casual conversation to a minimum during the drive and preferably, to restrict discussion to questions when they wanted clarification from the passenger during open and cued recall tasks (see below). In addition, the driver was instructed to drive in accordance with all applicable laws governing the use of roadways in the state of Maryland.
- Purpose: The overall aim of this research was to examine human brain and behavioral state changes during functional tasks performed in a naturalistic setting, here defined as a natural environment that is not characterized by typical laboratory or closed-course constraints. Examining the participants provided a unique opportunity to observe brain and body state dynamics in response to conditions involving authentic risks and consequences that are not accessible in any laboratory simulation or otherwise controlled environment. Driving on an open course involving public roadways with minimally-controlled traffic and environmental conditions incurs similar, though not the same, risks as those which individuals experience when they commute in an automobile. Though a task of increased naturalism and risk, roadway driving remains constrained enough to allow for application of established and replicable analytic techniques. That is, due to the nature of the legal and procedural requirements for safely and properly conducting the task, driving contains essential behavioral elements (e.g. real-life risk, behavioral stability, a multitude of repeatable events, such as braking) that preserve both ecological and experimental validity necessary to facilitate later interpretation of the captured data.

While they navigated the local roads in route to an interstate highway, an experimenter who was seated in the back seat pressed a button to begin the in-drive tasks of the experiment. First, the participants completed the Joint Auditory Verification Indicator (JAVI) task:

- Task: In this task, both the driver and passenger simultaneously listened to a pre-recorded health-focused (sleep or physical activity) podcast that was 2-3 minutes in duration, and was amplitude modulated at 40Hz. They were encouraged, but not required, to discuss or answer any questions about the content of this podcast during the drive.

- Purpose: The JAVI serves as a shared listening task where both participants engage the same content, just as during the *Memory Recall and Communication* tasks. In the JAVI task, however, the participants do not speak and therefore the underlying neural activity across the participants should be more alike than when the passenger talks and the driver listens. Here, this will be done to verify the validity of collection of joint EEG in this high-noise/high motion-artifact environment. That is, the increased similarity in task requirements for both participants provided a basis to estimate the level of neural synchrony that is measured during the inter-subject correlation (ISC) analysis, a core component of all three hypotheses. Further, the podcast was amplitude modulated to overcome the decreased signal-to-noise ratio in the real-world driving scenario by capitalizing on a research technique called steady-state evoked potentials (SSEPs). Also called *frequency-tagging*, this method uses different frequencies of visual flicker or amplitude modulation of auditory signals to target specific brain networks (for review, see Viallatte et al., 2010; Norcia et al., 2015), and SSEPs have also been highly successful in high noise EEG environments since they are narrow-band effects. Previous research has shown that a 40Hz modulation elicits a robust EEG oscillation (Galambos, Makeig, & Talmachoff, 1981), and it is outside the frequency ranges that dominate the task-dependent oscillations of interest for our experimental questions (e.g., delta, theta, alpha, beta). Thus, this EEG oscillation had a bandwidth and center frequency that is common among participants, providing a robust, low-level sensory signal to benchmark and verify ISC analysis. Finally, the content of the JAVI podcasts provided messaging stimuli regarding health behaviors that are believed to contribute significantly to inter-individual variability in social dynamics. Therefore, the specific content of the JAVI podcasts can also be paired with messaging protocols that occurred outside of the driving sessions as well as with belief questionnaires administered both before and after the driving sessions to examine the participants' responses to the social messaging (e.g. how their beliefs and behaviors changed as a function of the repeated messaging).

Once on the interstate, the staff member pressed a button to begin the next task, and this triggered a 5-minute timer and an instruction on the tablet resting in the passenger's lap to complete the open recall task:

- Task: Before the drive began, the passenger watched a set of video clips that communicated news items. In total, there were 32 individual news items, each of which had a humorous and a neutral variant. Trained actors narrated each news item once with each framing (humorous or neutral), yielding a total of (32 items x 2 actors x 2 framing) 128 possible clips. Any given participant only viewed a subset of 16 clips, half neutrally framed and half were humorously framed. Across the pair of dyad drives, each participant watched a separate, unique set of 16 news clips when they were the passenger. During the initial viewing, the humorously-framed and neutrally-framed items were shown in a random sequence mixing both types. During the highway portion of the drive, the task for the passenger was to communicate the content of the news clips to the driver in two complementary tasks: (1) an open recall and (2) a cued recall. During the open recall portion, the passenger was prompted to tell the driver as much as they could remember about any and all of the stimuli to which they were exposed. The driver could ask questions during this period, and the passenger was instructed to answer the questions as well as briefly engage any discussion that developed about the news items. This period lasted for a total of 5 minutes.

- Purpose: This task served as the primary method for addressing the scientific aims of the experiment. Verbal communication was chosen as a natural and safe means of exploring both intra- and inter-individual patterns of brain-body states that are related to individual and coordinated cognitive aspects of behavior. For the passenger, the communication task was set in a context of a dynamic emotional valence that is controlled by news item framing. We were interested in assessing the influence of this valence on the effectiveness of communication as well as the formation of trust in the driver. Within this context, the communication recall task enabled examination of the relationship between experimentally controlled factors of emotional valence (item framing), while the EEG analysis focused on correlation between brain signals as a proxy for shared representation of the discussion topic and communication effectiveness. Two different types of recall memory tasks were employed. The open recall portion provided an indication of which news items were encoded in the most salient fashion. For example, higher rates of recollection were expected for humorously framed as compared with neutrally framed items.

Following the completion of the Open Recall task, the participant began the cued recall task. This task is similar to the open recall task, with the exception that during cued recall the passenger received 16 text-based cues reminding them of each of the individual stimuli in a randomized order that was different from the order originally presented, and then they were instructed to discuss each with the driver for 1 minute. The purpose of the cued recall task was to allow experimental and analytic access to memories that require a small prompt, which engages slightly different cognitive mechanisms.

For cued recall, the tablet first provided a set of instructions then proceeded providing the first cue to talk about one of the news clips watched. Each cue consisted of a short text-based snippet of the news story. The passenger had 1 minute to describe each news clip, and the driver again listened and discussed or asked questions as appropriate. Halfway through this cued recall task, the task paused automatically while the driver exited the interstate to turn around and begin the return portion of the drive. During this time, the passenger was prompted on the tablet to complete their first rating of their trust in the driver using the Trust in Driver Scale (TiDS), as well as their second response to the SAM. They also indicated their perception of the current drive difficulty and the traffic density using the appropriate VASs (dVAS, tVAS).

To assess the passenger trust in the driver in a manner comparable to the trust in automation literature we used an adaptation of a system trustworthiness questionnaire that was successfully employed in a previous ARL experiment (#ARL 15-023; PI: Metcalfe). The TiDS is an adaptation of the trust ratings used in the foundational trust in automation research of Muir and Moray (1996); it comprises four visual analogue scales that allow the participant to rate their impression of the competence, predictability, reliability over time, and dependability of the driver based on his or her performance to the point in time when the instrument is administered. This questionnaire was administered only to the passenger, once at the halfway point of the drive and after completion of the full driving session. When analyzed, the participant's responses were coded as having a value between 0 and 100, representing all possible values between the far left and the far right ends of the line, respectively.

Once back on the interstate, the staff reinitiated the script, and the cued recall task continued. The passenger then proceeded to complete the remainder of the cued recall task. As the dyad exited the interstate and returned to the home location via local roads, they completed a second iteration of the JAVI task by passively listening to a second podcast about the benefits of physical exercise or sleep. After the audio was complete, the dyad completed the drive with minimal to no conversation as they returned to the home location.

Upon return to the home location, the experimenter immediately verbally administered the Motion Sickness Assessment Questionnaire (MSAQ) to document any change in motion sickness symptoms as a result of the drive. The MSAQ is a standard motion sickness screening tool and was administered to address an unlikely, but possible experience that occurs with driving environments involving attention to screens rather than out of the windshield (such as in motion simulation and indirect-vision driving experiments). Motion sickness symptoms may include, but are not limited to: nausea, cold sweating, pallor, and vomiting, and, if severe enough, these symptoms may impair task performance. Therefore, as both a safety precaution and as a means to objectively record any occurrences of motion sickness, even if minor, the MSAQ was administered to the passenger prior to the experiment (as a baseline), following the conclusion of the driving session, and in the event that a passenger experienced motion sickness symptoms during the driving session.

Next, the passenger used the tablet to complete the recognition memory task:

- Task: In this task, the participants were presented the first sentence from each of the 16 original news items that they previously viewed (OLD) and 16 lure news items that are similar, but not exactly the same as a previous statement (NEW). These 32 items were presented in a randomized sequence. Below each item was the question "Have you seen this item before?" and then they selected either OLD or NEW and indicated their confidence level in this judgment. Figure 2-41 presents an exemplar pair of old and new items as they were presented.
- Purpose: This task serves as the primary method to quantify the effect of emotional valence on memory performance. In the case of the passenger, the assessment is about the material that they directly viewed and then reinforced by talking aloud about it. For the driver, this serves as an assessment of how effectively the information was transmitted from the passenger, which is essential to ensuring that the task is about driver-passenger communication and not simply passenger memory. Since each participant answered these questions separately, this test provides an assessment of memory that is not intermixed with the dynamics of the dyad interaction; instead, it provided a means to quantify the success of dyad communication while simultaneously capturing any effects of the emotional valence (humor) manipulation on communication and memory. As such, it provided a critical parameter across all three of the

core hypotheses in this experiment.

Figure 2-42, Exemplar stimuli for the Recognition Memory Task with the original (OLD) item on the left and the lure (NEW) item on the right in which only the age has been changed.

While the passenger completed the recognition memory task, the driver used the Android phone to complete a second set of state scales (SAM, sVAS), the post-drive assessment of the passenger trustworthiness with the TWS, and provided their only subjective indications of their perception of the drive difficulty and overall traffic density (dVAS, tVAS). The driver also responded to the surveys regarding their attitudes and beliefs about sleep and physical activity (PABS, SHBS). After the passenger completed the recognition memory task the tablet was passed to the driver who then completed all of the above mentioned subjective instruments (dVAS, tVAS, sVAS, TWS, PABS, SHBS) as well as their second assessment of trust in the driver (TiDS).

Next the passenger departed the vehicle and the driver verbally completed the recognition memory task, reporting everything that they remember that the passenger told them about the news clips. Following the open recall task, the driver completed both the cued recall and recognition memory tasks as were previously completed by the passenger. The driver then exited the vehicle and the physiological sensors were removed.

Interim Period

There was a 2 – 5 week interim period between Drive A and Drive B, during which the participants provided the same data as described above for the baseline period (through their actigraph watches and online survey). The only difference was the addition of a reminder message about one of the facts presented in the podcast about the benefits of increasing physical activity or improving healthy sleep. The participants were then provided a brief yes or no answer to a question related to the fact. This question was in a text immediately following the message with the link to the online survey with the same questions as were answered during the baseline period (MRSI, DDSI, & DRQ).

Drive B

The procedure for Drive B was the same as for Drive A, but the passenger and driver reversed roles.

Follow-up Period

After Drive B, the participants continued to be monitored during a 2 – 5 week follow-up period, which was identical to the interim period, with the only difference being that the persuasive messaging facts were taken from the podcasts (on either physical activity or sleep habits) that the participants heard during the JAVI task of the *second* drive.

Close-out

At the conclusion of data collection for the entire experiment, all participants who agreed to complete the online Study Social Network Information Tool received an email with a link to the survey as well as to one final set of physical activity and sleep habit surveys (PABS, SHBS).

2.4.1.7.7 POC

Jason Metcalfe (jason.s.metcalfe2.civ@army.mil), ARL, Scientist
Jean Vettel (jean.m.vettel.civ@army.mil), ARL, Scientist
Javier Garcia (javier.o.garcia.civ@army.mil), ARL, Scientist
Rob Fernandez (rfernandez@dcscorp.com), ARL, Vehicle Instrumentation
Emily Falk (falk@asc.upenn.edu), Univ. of Penn., Scientist
Matt O'Donnell (matt.odonnell@asc.upenn.edu), Univ. of Penn., Scientist
Bret Kellihan (bkellihan@dcscorp.com), DCS Corp., Physiological Instrumentation
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Management
Michael Dunkel (mdunkel@dcscorp.com), DCS Corp., Data Engineer

2.4.2 CANCTA Publications

Vernon Lawhern, W. David Hairston, Kaleb McDowell, Marissa Westerfield, and Kay Robbins (2012), **Detection and Classification of Subject Generated Artifacts in EEG Signals Using Autoregressive Models**, *Journal of Neuroscience Methods*, 208, pp. 181-189.
DOI: 10.1016/j.jneumeth.2012.05.017
https://www.sciencedirect.com/science/article/pii/S0165027012001860

Anthony J. Reis, Jon Touryan, Jean Vettel, Kaleb McDowell, and W. David Hairston (2014), **A Comparison of Electroencephalography Signals Acquired from Conventional and Mobile Systems**, *Journal of Neuroscience and Neuroengineering*, Vol. 3, pp. 10-20.
DOI: 10.1166/jnsne.2014.1092
http://iopscience.iop.org/article/10.1088/1741-2560/11/4/046018/pdf

Belcher, M.A., Hwang, I., Bhattacharya, S., Hairston, W.D., & Metcalfe, J.S. (2022). **EEG-based prediction of driving events from passenger cognitive state using Morlet Wavelet and Evoked Responses.** *Transportation Engineering, 8,* 100107.
DOI: 10.1016/j.treng.2022.100107.

Dongrui Wu, Brent Lance, and Vernon Lawhern (2014) **Transfer Learning and Active Transfer Learning for Reducing Calibration Data in Single Trial Classification of Visually Evoked Potentials**, *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2801-2807, 2014
DOI: doi.org/10.1109/smc.2014.6974353
https://www.researchgate.net/publication/281526717_Transfer_Learning_and_Active_Transfer_Learning_for_Reducing_Calibration_Data_in_Single-Trial_Classification_of_Visually-Evoked_Potentials

Dongrui Wu, Vernon Lawhern, Brent J. Lance (2015) **Reducing Offline BCI Calibration Effort Using Weighted Adaptation Regularization with Source Domain Selection**, *2015 IEEE International Conference on Systems, Man, and Cybernetics Hong Kong*
DOI: 10.1109/SMC.2015.557
https://arxiv.org/pdf/1702.02897.pdf

Dongrui Wu (2017) **Online and Offline Domain Adaptation for Reducing BCI Calibration Effort**, *IEEE Transactions on Human-Machine Systems*, Volume: 47, Issue: 4, Aug. 2017
DOI: 10.1109/THMS.2016.2608931
https://www.researchgate.net/publication/281526533_Reducing_BCI_Calibration_Effort_in_RSVP_Tasks_Using_Online_Weighted_Adaptation_Regularization_with_Source_Domain_Selection

Kay Robbins, Kyung-min Su, and David Hairston (2017) **An 18-subject EEG data collection using a visual-oddball task, designed for benchmarking algorithms and headset performance comparisons**, *Data in Brief*, Volume 16, February 2018, Pages 227-230
DOI: 10.1016/j.dib.2017.11.032
https://www.sciencedirect.com/science/article/pii/S2352340917306285

S.M. Gordon, V. Lawhern, A.D. Passaro and K. McDowell (2015) **Informed decomposition of electroencephalographic data**. *Journal of neuroscience methods*, 256, pp.41-55.
DOI: 10.1016/j.jneumeth.2015.08.019
http://www.sciencedirect.com/science/article/pii/S0165027015003106

V.J. Lawhern, A.J. Solon, N.R. Waytowich, S.M. Gordon, C.P. Hung and B.J. Lance (2016) **EEGNet: A Compact Convolutional Network for EEG-based Brain-Computer Interfaces**, *Journal of Neural Engineering*, Volume 15, Number 5.
DOI: 10.1088/1741-2552/aace8c
https://arxiv.org/abs/1611.08024

S.M. Gordon, M. Jaswa, A.J. Solon and V.J. Lawhern (2017). **Real World BCI: Cross-Domain Learning and Practical Applications**. In *Proceedings of the 2017 ACM Workshop on An Application-oriented Approach to BCI out of the laboratory* (pp. 25-28). ACM.
DOI: 10.1145/3038439.3038444
https://www.researchgate.net/publication/314522756_Real_World_BCI_Cross-Domain_Learning_and_Practical_Applications

A.D. Passaro, J.M. Vettel, J. McDaniel, V. Lawhern, P.J. Franaszczuk and S.M. Gordon (2017) **A novel method linking neural connectivity to behavioral fluctuations: Behavior-Regressed Connectivity**. *Journal of neuroscience methods*, 279, pp.60-71.
DOI: 10.1016/j.jneumeth.2017.01.010
https://doi.org/10.1016/j.jneumeth.2017.01.010

A. Solon, S. Gordon, V. Lawhern and B. Lance (2017) **A GENERALIZED DEEP LEARNING FRAMEWORK FOR CROSS-DOMAIN LEARNING IN BRAIN COMPUTER INTERFACES**. In *The First Biannual Neuroadaptive Technology Conference* (p. 34).
http://neuroadaptive.org/conference/2017/programme

S.M. Gordon, A.J. Solon, M.W. Nonte, V.J. Lawhern, **Informed Decomposition to Improve Feature Extraction for EEG-Based BCI**, submitted to *IEEE Transaction on Biomedical Engineering*

Anne-Marie Brouwer, Maarten A. Hogervorst, Bob Oudejans, Anthony J. Ries, Jonathan Touryan (2017) **EEG and Eye Tracking Signatures of Target Encoding during Structured Visual Search**, *Frontiers in Human Neuroscience* 11:264
DOI: doi.org/10.3389/fnhum.2017.00264
http://journal.frontiersin.org/article/10.3389/fnhum.2017.00264/full

Anne-Marie Brouwer, Anne smelting, Matthew Jaswa, Oded Flascher, Laurens Krol, Thorsten Zander (2017) **Physiological Effects of Adaptive Cruise Control Behaviour in Real Driving**, BCIforeal'17 (Limassol, Cyprus, 13 MAR 2017), *Proceedings of the 2017 ACM Workshop on An Application-oriented Approach to BCI out of the laboratory* Pages 15-19
DOI: 10.1145/3038439.3038441
http://dl.acm.org/citation.cfm?doid=3038439.3038441

L.R. Krol, T.O. Zander, M. Jaswa, O. Flascher, A. Snelting, A-M. Brouwer (2017) **Online-Capable Cleaning of Highly Artefactual EEG Data Recorded During Real Driving**, *Proceedings of the 7th Graz Brain-Computer Interface Conference 2017*
DOI: 10.3217/978-3-85125-533-1-47

http://castor.tugraz.at/doku/BCIMeeting2017/47.pdf

Lin, C. T., Prasad, M., Gupta, A., Ma, Z. Y., Lin, J. W., Singh, A. K., Machado, A. M. C., Appaji, A., and King, J. T. (2018) **Voice Navigation Effects on Real-World Lane Change Driving Analysis Using an Electroencephalogram**, *IEEE Access* 6: 26483-26492.
DOI: 10.1109/ACCESS.2018.2820161
https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8331266

Wang, Y. K., Jung, T. P., and Lin, C. T. (2018) **Theta and Alpha Oscillations in Attentional Interaction during Distracted Driving**, *Frontiers in Behavioral Neuroscience*, 12:3
DOI: 10.3389/fnbeh.2018.00003
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5811509/

Lin, C. T., King, J. T., Gupta, A., Ma, Z., Fan, J. W., Appaji, A., Machado, A. M. C., and Prasad, M. (2017) **The Influence of Acute Stress on Brain Dynamics During Task Switching Activities**, *IEEE Access*, Vol. 6, Issue 1, pp. 3249-3255, Dec. 2017.
DOI: 10.1109/ACCESS.2017.2787673
https://ieeexplore.ieee.org/document/8240583/

Wang, Y.T.; Cheng, C.K.; Huang, K.C.; Lin, C.T. (2012) **Cell-Phone Based Drowsiness Monitoring and Management System**. In Proceedings of the IEEE International. Conference. Biomedical Circuits and Systems. Hsinchu, Taiwan, 28–30 November 2012; pp. 200–203.
DOI: 10.1109/BioCAS.2012.6418462
https://ieeexplore.ieee.org/document/6418462/

Kerick, S., Chuang, C.-H., King, J.-T., Jung, T.-P., Brooks, J., Files, B., McDowell, K. and Lin, C.-T. (under review). Inter- and intra-individual variations in sleep, subjective fatigue, and vigilance task.

Fonseca, A., Kerick, S., King, J. T., Lin, C. T. & Jung, T. P. (under review). Brain network changes in fatigued drivers: a longitudinal study in a real-world environment based on nonlinear EEG analysis and actigraphy data.

Do, T. T. N., Wang, Y. K., King, J. T. & Lin, C. T. (under review). Increase brain information flow when workload increases but not with high fatigue state.

Huang, K. C., Hsieh, C. Y., Chuang, C. H., King, J. T. and Lin, C. T. (under review). The effects of different chronic fatigue levels on brain-behavior relationships in driving.

Huang, K. C., Wang, Y. K., Chuang, C. H., King, J. T. and Lin, C. T. (under review). EEG information transfer changes on fatigue drivers.

Lin, C. T., Nascimben, M., King, J. T., Wang, Y. K., Jung, T. P. and Kerick, S. (under review). EEG and HRV entropy task related factors under different fatigue levels.

Vettel, J.M, Lauharatanahirun, N., Wasylyshyn, N., Roy, H., Fernandez, R., Cooper, N., Paul, A., O'Donnell, M.B., Johnson, T., Metcalfe, J., Falk, E.B., & Garcia, J.O. (2018). **Translating driving research from simulation to interstatedriving with realistic traffic and passenger interactions.** In D. Cassenti (Ed.). *Advances in Human Factors in Simulation and Modeling. AHFE 2018. Advances in Intelligent Systems and Computing, 780*.
DOI: *10.1007/978-3- 319-94223-0_12*.

2.4.3 CANCTA Datasets

The following list of datasets represents partial accounting of data collected under the CANCTA program; i.e., those that are in, or being processed for the SANDR.

2.4.3.1 ARL_EEGCS_VEP Dataset

A total of 18 datasets were collected across 18 recording sessions, from 18 unique subjects recruited from within ARL HRED at Aberdeen Proving Ground.

The total size of the EEG files is 3.26 GB, representing 2.82 hours of EEG recording.

Note: The dataset represents only the BioSemi EEG system and VEP task combination.

2.4.3.2 DCS_CANCTA_FT Dataset

A total of 13 datasets were collected across 13 recording sessions, from 13 unique subjects recruited from within ARL HRED at Aberdeen Proving Ground.

The total size of the EEG files is 46.3 GB, representing 18.21 hours of EEG recording.

2.4.3.3 DCS_CANCTA_ODE Dataset

A total of 67 datasets were collected across 17 recording sessions, from 17 unique subjects recruited from within ARL HRED at Aberdeen Proving Ground. Each subject performed 4 subtasks within the ODE task, resulting in an EEG file. There was 1 dataset that was not usable.

The total size of the EEG files is 15.3 GB, representing 18.74 hours of EEG recording.

2.4.3.4 NCTU_CANCTA_RWN_VDE Dataset

A total of 855 datasets were collected across 855 recording sessions, from 17 unique subjects recruited from within the NCTU college student body. Each subject performed 5 task phases within the data collection session, resulting in an EEG file. There was 1 dataset that was not usable.

The total size of the EEG files is 129.8 GB, representing 207.62 hours of EEG recording.

2.4.3.5 TNO_CANCTA_FLERP Dataset

A total of 42 datasets were collected across 21 recording sessions, from 21 unique subjects recruited by TNO (Netherlands). Each subject performed 2 subtasks within the FLERP task, resulting in an EEG file.

The total size of the EEG files is 16.56 GB, representing 23.02 hours of EEG recording.

2.4.3.6 TNO_CANCTA_ACC Dataset

A total of 45 datasets were collected across 15 recording sessions, from 15 unique subjects recruited by TNO (Netherlands). Each subject performed 3 subtasks within the ACC task, resulting in an EEG file.

The total size of the EEG files is 33.20 GB, representing 22.50 hours of EEG recording.

2.4.3.7 RWN-VDEDP Dataset

A total of 43 duel-subject (driver and passenger) datasets were collected across 43 recording sessions, from 44 unique subjects paired into 22 dyads. The subjects were recruited from ARL HRED at Aberdeen Proving Ground.

The total size of the EEG files is 33.85 GB, representing 129.77 hours of EEG recording.

2.5 Autonomy Research Pilot Initiative (ARPI)

Vehicle survivability is an important issue in today's military. One of the most critical influences on survivability is the performance of human operators – especially as it degrades with time on task, fatigue, divided attention, and other factors. In response to current challenges, significant DoD research and development efforts have focused on integrating autonomous vehicle technologies as one strategy for mitigating the impact of human error on overall system performance. However, simply implementing autonomy without having a clear scheme for *integrating* human and automated task inputs has led to relatively poor system performance and thus low user acceptance.

The degree to which human operators accept and appropriately use autonomous technologies is heavily influenced by issues of trust and confidence. That is, as people gain confidence in the reliability, robustness, and safety of autonomous technologies, they develop sufficient trust to willingly share important decision and/or control authority with such systems. It has been well-documented that human trust in automation (TiA) is considered a major determinant of acceptance and use of these advanced technologies (Dzindolet et al., 2003; Lee & Moray, 1992; Lee & See, 2004; Muir, 1989, 1994; Parasuraman & Riley, 1997). More important than achieving a certain level of trust, however, is to find an appropriate match between the relative capabilities of the technology and the operator's trust, at least inasmuch as trust affects system use, misuse, or disuse (Parasuraman & Riley, 1997; McBride & Morgan, 2010; McGuirl & Sarter, 2006; Merritt et al., 2014). Therefore, an important goal for systems designers is to find means and mechanisms for ***calibrating*** human TiA *to **elicit the desired interaction behavior*** with the autonomy given the nature of the ongoing, and dynamic, task context (Cai & Lin, 2010; Inagaki, 2005; Jamieson & Vicente, 2005; Lee & See, 2004; Parasuraman, 2000). Of course, in order to develop a system for active, online calibration – or, effective management – of TiA, one first must have reliable and objective indicators of how and when it changes. Moreover, if such indicators are to prove useful, then they must also be robust within and between individuals as well as across task, time, and technology. In the ARL_ARPI_TX20 experiment we ask whether, through monitoring an operator's state and behavior, objective metrics can be developed that enable effective inference of his or her TiA. If the objective, real-time inference of TiA can be demonstrated as valid, reliable, and robust, then systems could be designed to continuously monitor TiA in order to maximize the effectiveness (while minimizing the risk) of critical human-autonomy shared tasks. In the follow-up ARL_ARPI_TX22 experiment we asked if we could (a) predict instances where a subject is likely to either yield control to ("handoff) or arrest control from ("takeover") an automation in real-time and (b) if we can use that information with appropriate feedback to influence the likelihood of one behavior over another given previously observed response patterns indicating which agent would be the optimal performer given the current circumstance.

In addition to robust and reliable inference of TiA, the notion of its real-time management implies a degree of *prediction*. That is, predictive mitigation strategies are highly preferred to reactive compensation for negative consequences already incurred – especially as consequences may involve extremely expensive assets such as unmanned ground and air vehicles or even semi-autonomous vehicles, which carry more serious risks associated with the safety of human occupants. Indeed, some have taken the approach of predicting future system performance as a means of informing and enhancing operator trust (e.g. Cai & Lin, 2010; Gu et al., 2014). However, while somewhat effective when used appropriately, this type of method is rather indirect and rests on assumptions that the operator will always use the information provided in appropriate and predictable ways. An alternative is to measure the operator directly in real time and then use an *a priori* understanding of indicators of trust-related behaviors (disuse, misuse) to predict when system mitigations may be necessary. Yet, owing to heavy use of subjective self-report and/or performance based evaluations, the accuracy and bandwidth of prospective information available about operator state have been typically limited. Therefore, significant progress has yet to be made in the development of useful operator-centered estimators and predictors that could enable real time trust management.

Indeed, considerable research on interpersonal trust has revealed measurable patterns of physiological change that correlate significantly with changing levels of subjective trust and trust-based decision making (c.f. Borum, 2010 for an extensive review). However, to date, trust measurement in the context of human-autonomy interaction still exists largely as a subjective endeavor (Dzindolet et al., 2003; Lyons et al., 2011; Moray, Inagaki, & Itoh, 2000). Encouragingly, methods and measures from psychophysiology have been used to gain insights into other aspects of human-autonomy interaction that are likely to be related to variations in TiA. For instance, despite the reality that few studies have used eye-tracking expressly for trust inference, it has been used to gain insight into constructs such as visual attention, mental workload, and automation complacency (de Greef et al., 2009; Metzger et al., 2000; Rovira & Parasuraman, 2010), each of which have logical and/or empirical connections with TiA. Similarly, electrophysiological measurements, such as electroencephalogram (EEG), electrocardiogram (ECG), or electrodermal activity (EDA) monitoring, have not been applied directly to the inference of trust in automation as a cognitive state. However, as with eye tracking, psycho- and electro-physiological methods have been applied to examine constructs, like task engagement and technology acceptance, that have expected associations with TiA (Fairclough & Venables, 2006; Moridis et al., 2012). Moreover, fMRI research has demonstrated differential brain activations for building (paracingulate cortex) versus maintaining (ventral tegmental area) interpersonal trust (Krueger et al. (2007); importantly, the association of trust maintenance with the ventral tegmental area also leads to a potential link to online evaluation of expected risk versus reward. In the ARL_ARPI_TX22 experiment we hypothesize that such results will extend to TiA-based decision making as a special case of value-based decision-making (Rangel et al, 2008) and will thus be reflected in changing EEG as well as associated peripheral physiology (ECG, EDA).

Summarily, the research conducted under the ARPI program is aimed at enhancing human-autonomy interactions by facilitating the eventual real-time influencing or biasing of behaviors that are thought to be based on TiA (i.e. handoff and takeover decisions). The primary focus is to assess and further develop psychophysiology-based prediction of changes in trust and/or trust-related behaviors for the specific purpose of providing feedback that will influence more optimal usage of the driving automation.

2.5.1 ARPI Program Summary

Under the Autonomy Research Pilot Initiative funded by the Office of the Secretary of Defense, research in the Real-World Soldier Quantification branch of ARL-HRED, in collaboration with the Army Ground Vehicle Systems Center (GVSC, formerly the Tank and Automotive Research, Development and Engineering Center - TARDEC, at the Detroit Arsenal, Warren, MI), the Air Force Research Laboratory (AFRL; Wright-Patterson AFB, Dayton, OH), and the US Navy Space and Naval Warfare Systems Command (SPAWAR; San Diego, CA), aims to develop a broad framework for enhancing the integration of humans and autonomy in the use of advanced technologies, including military vehicle-based systems. Measurement and Inference of Trust in Driving Autonomy as a function of Reliability and Workload (ARL_ARPI_TX20)

2.5.1.1 Measurement and Inference of Trust in Driving Autonomy as a function of Reliability and Workload (ARL_ARPI_TX20)

This research was conducted to develop and validate methods for monitoring and predicting varying degrees of trust in automation (TiA) using both physiological and behavioral metrics characterizing real-time human-automation interactions. The overarching goal of this research was to develop and validate methods for measuring and drawing inferences about TiA, either directly or indirectly through correlated constructs. In particular, we examined operator trust in vehicle automation as it is reflected in changes observed in subjective reports as well as behavioral and physiological state variables during the execution of a shared human-autonomy driving task. The stated aims underlying this goal included:

Aim #1: To develop and experimentally validate metrics (dependent variables) that index changes in TiA. Rather than focusing on single-modality metrics, we will record and explore the patterns of correlation and co-variance among a variety of psychophysiological and behavioral variables and focus particularly on metrics that predict decisions around sharing vehicle control with the autonomy in each condition. State measures will be derived from EEG, EOG (electrooculography), ECG, EDA, and gaze position tracking as well as the subject vehicle control behaviors.

Aim #2: To develop an understanding of factors (independent variables and covariates) that influence the subject's TiA. Whereas the *Aim #1* targets the identification of metrics, or groups of metrics, that reliably predict trust-based decision-making, here we seek to gain insight as to which factors influence the likelihood and directionality of those same trust-based decisions. Such factors will include real-time tracking of variables such as task load, collision risk, and recent performance history or trending changes in success rate.

2.5.1.1.1 Protocol

The research protocol information for this study is as follows:

> Organization: ARL
> Protocol Number: ARL-15-023
> Protocol Name: "Measurement and Inference of Trust in Driving Autonomy as a function of Reliability and Workload"
> Contract: W911NF-10-D-0002

2.5.1.1.2 Location

This study was conducted at GVSC's Ground Vehicle Simulation Laboratory (GVSL) in Warren, MI, using the Ride-Motion Simulator (RMS).

2.5.1.1.3 Subjects

A total of 24 subjects participated in this study. Each subject completed all tasks in a single day. A total of 119 data recordings were produced, including 16 for the PRACTB task, 23 where only manual control of the vehicle was available, and 80 where the different levels of autonomous vehicle control were available.

2.5.1.1.4 Apparatus

The following items were used during the experiments. All equipment is located within the GVSL at GVSC.

Ride Motion Simulator (RMS): The RMS is a 6 degree-of freedom (DOF) motion-based simulator that is capable of reproducing the dynamics of military ground vehicles over a vast array of terrains that are commonly experienced by current force vehicles. As shown in Figure 2-42, the RMS is composed of a platform that is mounted on a hexapod design of hydraulic actuators, the coordinated motion of which produces precise resultant motions through 3 cardinal linear (longitudinal, lateral, and vertical) and 3 rotational (yaw, pitch, roll) dimensions. The RMS supports a re-configurable cab that allows the experimenter to customize the driver controls, incorporate rugged display monitors, seats, and data collection devices (e.g. eye tracking and EEG). Supported by the Real-Time Simulation Framework, this environment allows the user to be immersed into a synthetic terrain of various types and experience realistic ground vehicle dynamics. It has integrated motion, audio, and visual systems for high fidelity real-time simulation.

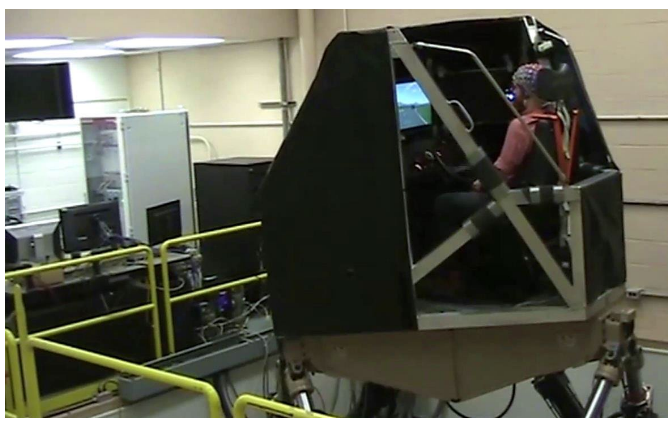


Figure 2-43. Ride Motion Simulator

Crewstation Cab: The simulation environment was constructed to present participants with visual, motion and audio cuing in order to recreate a realistic driving experience. For this experiment, the cab was configured as a single-occupant crew station with 3 flat panel displays, thus allowing creation of central and peripheral visual flow patterns. The simulator cab comprised a vehicle seat with a 4-point harness, a steering buck with wheel and standard gauges (speedometer, tachometer, etc.), throttle and brake controls on the floor, and custom controllers for additional task inputs. Audio cuing was limited to presenting the participant with the experimenter's voice (during instruction and other communication periods), the vehicle's sounds (engine noise correlated to engine RPM), and driving-related environmental sounds (e.g. wind gusts, passing vehicles).

Electrophysiological Data Recording: Electrophysiological data was recorded using a commercially-available Biosemi ActiveTwo system. Electroencephalographic (EEG) data was collected using an electrode cap with 64 active surface electrodes. The ActiveTwo system uses sintered Ag/AgCl electrodes held in place by a snug elastic fabric cap. This system does not require conventional skin cleaning and preparation, such as abrasion, but does require gel; hypoallergenic, bacteriostatic gel was used to establish conductive contact. The procedure for placing the electrodes was standard, minimally discomforting, and has been extensively used in both clinical practice and research. The configuration of the electrophysiological recording system included 5 bipolar pairs: vertical (VEOG) and horizontal (HEOG) electrooculogram, electrocardiogram (ECG), 2 natively referenced channels (left and right mastoid process), and one EDA channel. The VEOG pair was centered superior and inferior to the left eye. The HEOG pair was placed along the outer canthus of each eye. The ECG pair was placed on the notch atop the sternum and clavicle of the participant. The bipolar EDA was placed on the palmar surface of the distal phalanges or the thenar and hypothenar eminences of the left or right hand (whichever hand is not the one that the participant identifies as her/his primary steering hand).

Head and Eye Tracking: Minimally calibrated eye gaze and head movement data was acquired using a commercially-available eye tracking system (Smart Eye AB) comprising infrared emitters and four computer controlled optical zoom cameras that allow 6 degree-of-freedom head tracking, 2 degree-of-freedom gaze tracking, and eye blink detection and pupil diameter monitoring.

Video/Audio Recording System: Video data was captured using 2 video cameras mounted on the inside of RMS cab, one aimed at the driving display and the other at the subject's face. All video data was used for real-time subject monitoring and is available to facilitate interpretation of quantitative data. Audio data was recorded in order to capture dialogue between the participant and investigators.

Standard PCs, Monitors, and Input Devices: This project required connecting an array of standard PCs as a local network to enable data synchronization and storage. The RMS control and simulation framework used a small cluster of high-speed computational and network assets, with software that was optimized for real-time operations with low latency, yet able to preserve physics-based fidelity of vehicle and terrain models. Additional PCs were used for audio/video recording, eye tracking, surface electrophysiology (EEG/EOG/ECG/EDA), and for the presentation of task stimuli. Responses to pedestrian appearances were made using a commercially-available game controller (the Nintendo® Wii™ remote).

2.5.1.1.5 Demographics

Demographic information collected for this dataset includes subject ID, age, and dominant hand, however, this data was not processed in the current version of the data. It's available in the raw data only.

2.5.1.1.6 ARL_ARPI_TX20 Tasks

Participants performed a Physiological Response and Artifact Characterization Test Battery (PRACTB) control task and the main task, Leader-Follower Driving (LFD).

2.5.1.1.6.1 PRACTB Task

PRACTB is a series of simple laboratory tasks, most of which are standard for experiments that employ psychophysiological techniques (i.e. simple baseline, simple reaction time). It was used to provide baseline behavioral (eye, head, limb movement) and psycho-physiological (EEG, ECG, etc.) data sufficient to allow for experimental control.

For the ARL_ARPI_TX20 experiment, five PRACTB tasks were chosen, each intended to address specific needs for baseline data, sensor cross-validation, and artifact detection and extraction. Figure 2-43 shows the display presented to the participant for initiation of the tasks.

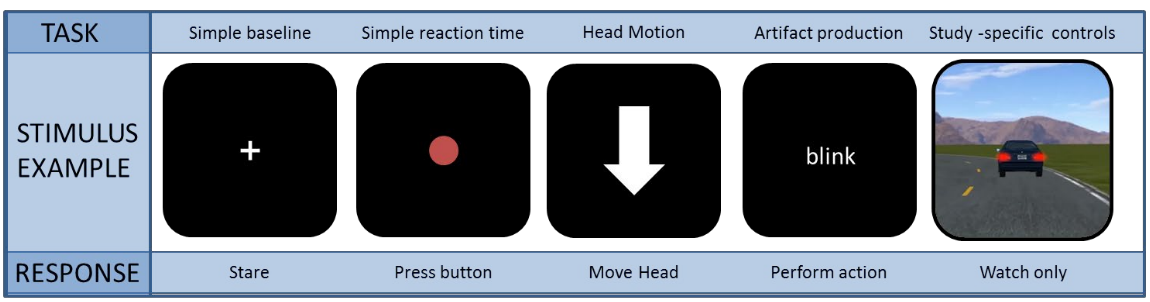


Figure 2-44. Control Screen for the Five PRACTB Tasks

1. *Simple baseline* – Data acquired from this task enabled characterization of the subject's baseline neurocognitive state before experimentation. The subject was asked to sit still and focus their gaze on a cross superimposed and centered on a uniform background. The baseline video is 2 minutes long.
2. *Simple reaction time* – Data from this task provided individual estimates of the subject's base visual reaction time using the thumb and forefinger of the hand to be used for the pedestrian task. The subject was asked to perform a simple reaction time task, composed of the presentation of a single red dot followed by a button press. Subjects were instructed to respond as quickly as possible. Stimuli were presented for a maximum of 1.5 seconds; each stimulus disappeared as soon as the subject responded or the presentation interval expired. To match the presentation rate of the pedestrian task (see below), the inter-stimulus interval was chosen from a uniform

random distribution between 4.5 and 6.5 seconds. There were 20 consecutive repetitions for each finger (thumb and index finger), with a 5 second pause when switching fingers.

3. *Head Motion* – The purpose of this task was to generate a set of head movement artifacts for later characterization, detection, and removal from EEG. For this task, the stimuli included a psuedo-randomized series of arrows that indicate a direction in which the subject is to rotate their head (left, right, up, down). Subjects were asked to move their head 'comfortably and naturally' and then return to a neutral position. The stimulus appeared for 1 second and inter-stimulus intervals were 2 seconds. There were 5 repetitions of each direction.
4. *Artifact Production* – The purpose of this task was to provide data on artifacts that were not already embedded in the previous tasks in a controlled fashion. Subjects were asked to produce a series of actions including: eye blink, eyebrow raise, shrug, and jaw clench. The actions were cued by a text description centered on the computer screen. There were 20 consecutive repetitions for each action. Each stimulus word appeared for 500 milliseconds and inter-stimulus intervals were 2 seconds.
5. *Study Specific Baselines* – The purpose of this last task was to provide baseline physiological data for later analyses of the leader-follower driving tasks. In this study-specific baseline, the subjects rode a vehicle through a small portion of the experimental environment, which contained the lead vehicle, but excluded other ambient traffic, pedestrian stimuli, and other screen feedback. The physical simulation apparatus (the RMS) was turned on throughout this baseline condition. Subjects were instructed to passively watch the out-the-window view of the simulated vehicle without taking any actions.

2.5.1.1.6.2 LFD Task

Subjects performed a set of tasks that represented challenges typically associated with driving. Some aspects of the driving tasks were, at times, performed jointly with a pseudo-autonomy that had varying degrees of performance reliability. Task objectives included lane position maintenance, maintaining a prescribed following distance behind a lead vehicle, and monitoring for and responding to pedestrians that appeared on the side of the road; each of these objectives had specific associated challenges, to be described below. The pedestrian monitoring task provided a persistent, high-consequence demand throughout each condition that encouraged the subjects to use the automation as much as possible, thus pressuring subjects to continuously make active decisions whether to trust the autonomy.

For each approximate 12-minute condition, subjects were instructed to drive a simulated vehicle (denoted "ownship") one full lap around the two-lane course while maintaining a distance within a prescribed range (5 – 20 m) behind the lead vehicle. Figure 2-44 illustrates the complete driving course.

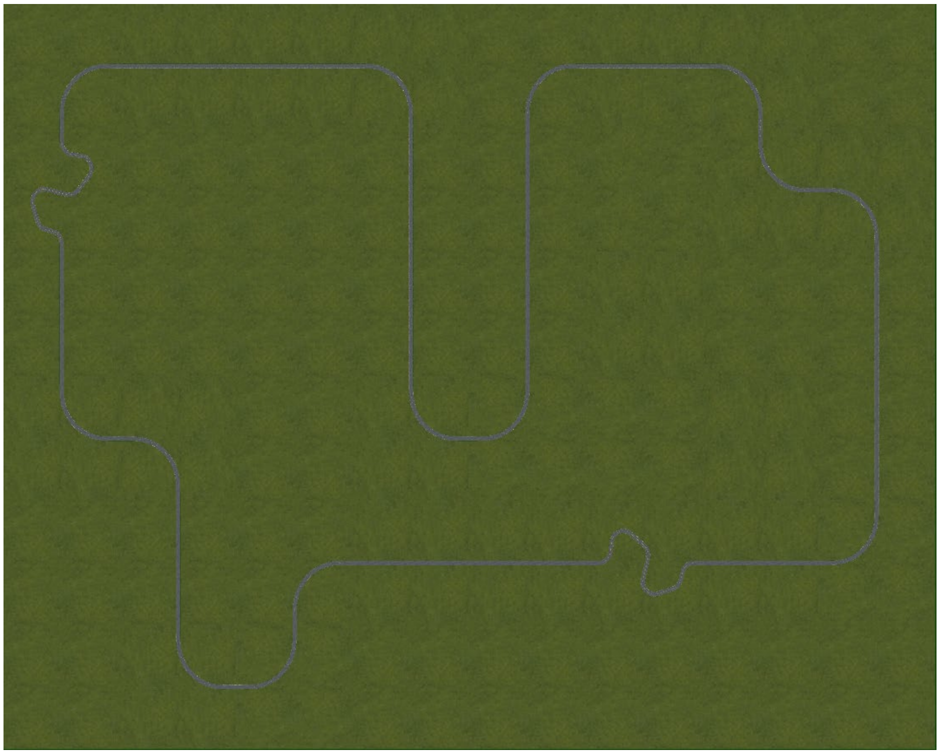


Figure 2-45. Driving Course for ARL_ARPI_TX20 Experiments

Participants were given screen feedback indicating whether they were in the required range. The screen display presented to the participant is shown in Figure 2-45. The green text string in the bottom center of Figure 2-45 indicates the ownship is within the required following range. The text string turned red when indicating the vehicle was outside of the required range limits. Speed was determined by the lead vehicle and road characteristics (i.e. curves) and, by design, the lead vehicle maintained an average speed of 25 mph, with perturbations and curves leading to speeds temporarily varying between 15 and 30 mph. Moreover, participants were instructed that they must not deviate outside of their lane and that all collisions should be avoided. Under different conditions, a pseudo-autonomy (here forward called “the autonomy”) was controlling either the vehicle speed (speed control, SC) or both speed and steering (full control, FC). Neither of the autonomies were capable of collision avoidance. For conditions in which the autonomy was enabled, subjects had the option to enable (using a dedicated on-toggle switch at their left foot) or disable the autonomy (by touching a control – throttle or brake – pedal with their right foot) at any moment. An indicator light on the screen provided constant status of the autonomy, with green meaning “enabled” and red meaning “disabled” (Figure 2-45, top right of main display window). When the autonomy was enabled, the subject’s vehicle controls did not affect the simulated motion, depending on the condition. Specifically, the steering wheel continued to function when the autonomy was enabled in the SC condition because the subject was expected to manage heading while the autonomy managed speed; neither the steering wheel nor the pedals functioned when the autonomy was enabled in the FC condition. In addition to autonomy condition, there were two automation reliability conditions (defined as described below), resulting in a 2 (autonomy type: SC, FC) × 2 (autonomy reliability: low, high) design with a manual driving baseline.

Figure 2-46. ARL_ARPI_TX20 LFD Screen Presentation

Figure 2-46 provides a summary of the LFD paradigm. Subjects drive a vehicle ('ownship', green) while following a lead vehicle (orange). Subjects are instructed to maintain a particular following distance as well as consistent lane position. In some conditions, the subject shares vehicle control with an auto-driver of varying reliability (low and high) as represented by the distributions labeled σ_L and σ_F. Environmental challenges are presented in the form of ambient traffic (f_T), perturbations to ownship heading (P_H) and lead vehicle speed (P_S), and static and dynamic pedestrians which appear with a specified frequency (f_p) and probability structure (probability of dynamic pedestrian, ρ_d).

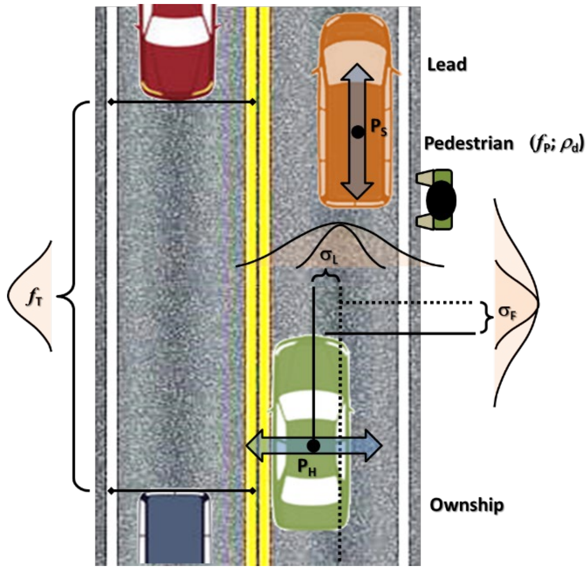


Figure 2-47. Summary of the LFD paradigm for ARL_ARPI_TX20

The reliability of each autonomy was defined according to the distributions shown in Figure 2-46 and labeled as σ_F and σ_L; these distributions define the amount of variability in maintenance of following distance and lane position respectively. For each performance variable in Figure 2-46, the desired lane position/ following distance is shown with the dotted line and the current position of the ownship is indicated with the solid line; the distributions represent the variance of the deviations around the dotted lines (the difference between the dotted and solid lines). For each performance variable, the wider distribution represents the low reliability condition and the narrow distribution represents the high reliability condition. Moreover, during the driving task, subjects experienced vehicle perturbations in both lateral and longitudinal directions. Lateral, or heading, perturbations (P_H) were presented as gust-like forces acting to push the vehicle out of its lane. Longitudinal, or speed, perturbations (P_S) involved changes in the speed of the lead vehicle through sudden positive (throttle) or negative (brake) accelerations. These perturbations occurred, on average, once every 15 ± 5 sec (normally distributed), resulting in approximately 4-5 perturbations per minute and approximately 45 - 55 per condition; no perturbations overlapped in time, meaning that a lateral and longitudinal perturbation did not occur simultaneously. The perturbations were semi-randomly selected as to type (lateral/longitudinal) and direction (positive or negative along relevant axis). As with lane position and following distance, the reliability condition impacted the ability of the autonomy to respond to these perturbations (low reliability took longer to return to the desired performance as compared with high-reliability). During the task, any collisions with other vehicles or pedestrians did *not* have a motion consequence (the RMS shutter as if driving over a “rumble strip” on the road), a visual collision indicator appeared on the screen.

Two factors influenced the sequence of perturbations in a deterministic fashion. First, our prior experience with simulated driving has shown that participants have difficulty resolving and responding to multiple sequential perturbations in a longitudinal direction; this is because it can be difficult to distinguish the end of one speed perturbation and the start of another. Therefore, there were at least 2 lateral perturbations for each longitudinal perturbation. Second, in order to empirically differentiate task difficulty from trust effects, vehicle perturbations in the high-reliability SC conditions had a different allocation than the other conditions. That is, in the high automation reliability condition, we expected that subjects would allow the vehicle to control responses to nearly all of the longitudinal perturbations and, as a result, they would respond to fewer total perturbations than in the low reliability condition and presumably have lower workload. Therefore, to balance workload/task difficulty across these two (SC) conditions the balance of lateral to longitudinal perturbations was a 3:1 ratio rather than 2:1 in all other conditions (this value was selected based on pilot testing). In the two full autonomy conditions, no effort was made to balance workload, thus allowing investigation of changes or interactions associated with automation reliability and workload/task difficulty.

In order to encourage as much use of the autonomy as possible, subjects also performed a task for which there were no possible responses from the autonomy. This task was designed to be relevant and similar to typical driving in that it involved reacting to pedestrian entities, some of which stepped directly into the pathway of the vehicle. Although this task was unlike real-world driving in that the pedestrians "popped" into of the scenario, it did reflect more infrequently occurring natural events such as a bicyclist or runner that suddenly emerges from behind a parked car. The pedestrian task was chosen to be presented in this fashion in order to facilitate analyses that have experimental analogs, such as the classic "oddball" task (see Garrido et al., 2009 for review and Squires et al., 1975 for original publication), within standard laboratories that study cognitive function. Throughout all driving conditions, pedestrians dynamically appeared on either the left or right shoulder of the road with a 50% probability; 15% of the pedestrians (the rare or "oddball stimulus") immediately began walking into the path of the vehicle when they appeared. Subjects were asked to respond to *all* pedestrians using a button press on a commercially-available game controller (the Nintendo® Wii™ remote), with static and dynamic pedestrians assigned to the thumb and forefinger of the same hand, respectively. This was kept the same across all participants because, based on pilot testing, it was decided that the differential control between thumb and forefinger was enough to preserve across subjects rather than counterbalance and potentially wash out effects. The response device was held in the same, subject-selected hand throughout all conditions of the experiment and the other was used for control of the steering wheel. Pedestrians appeared at approximately double the rate of vehicle perturbations; specifically, a new pedestrian appeared every 6 ± 1 sec. Once they appeared, a pedestrian remained in the scenario until the vehicle drove past it (approximately 1.5 sec) or until the participant respondeds with a correct button press, in which case the pedestrian simply disappeared. It was possible to "collide" with a pedestrian, however, these collisions were not represented with any physical consequences; the vehicle simply drove through the pedestrian image and a large, salient error signal was displayed on the screen. Moreover, if the initial button press response of the subject is incorrect, they had the opportunity to clear the pedestrian out of the way with a successive correct button press. Therefore, an incorrect response to a pedestrian did not necessarily mean that the ownship would collide with that pedestrian. Figure 2-47 shows the display following a pedestrian collision (errors for an incorrect button push and an "out of range violation" have also occurred per the bottom two red text strings).

Figure 2-48. Display Following a Collision with a Pedestrian for ARL_ARPI_TX20

Performance was quantified according to the following scoring structure. Participants began each LFD task with 500 pts. For the most part, points were not gained, but only lost as a consequence of performance errors. Score feedback was provided in real time using minimally-disruptive, centrally-presented pop-up values as well as a persistent score counter at the top left corner of the central display (Figure 2-45, light blue text string). To highly incentivize attention to the pedestrian task, each collision with a pedestrian resulted in a significant loss in points (-100 pts per collision); categorization errors through incorrect button press responses only incurred a small penalty (-5 pts). Collisions with other vehicles resulted in a moderate penalty (-50 pts per vehicle collision) and driving errors leading to lane departures or violations of the following distance threshold, incurred the lowest penalties (-2 pts per 3 seconds spent in driving noncompliance). To discourage subjects from completely giving up on the task if all points were lost, a small bonus (10 points) was given at the completion of each of 9 zones throughout the driving course, which resulted in a 100 point accumulation throughout each scenario. The final (total) score for the experiment determined the amount of "bonus" pay received by the participants. The point accumulation resulted in all participants being guaranteed at least a 20% bonus on their performance.

2.5.1.1.7 POC

Jason Metcalfe (jason.s.metcalfe2.civ@army.mil), ARL, Primary Investigator
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer
Michael Dunkel (mdunkel@dcscorp.com), DCS Corp., Data Engineer

2.5.1.2 Predicting and Influencing Trust-Based Decision Making during Human Interaction with Driving Automation (ARL_ARPI_TX22)

We studied strategies for influencing or biasing trust-based behaviors, and we used both physiological and behavioral metrics to characterize real-time human-automation interactions. The overarching goal of this research was to use the inferences about TiA that were developed in ARL_ARPI_TX20 to enable real-time prediction of behaviors based on TiA (automation handoff and takeover events) and to use such predictions to influence contextually appropriate automation use. The aims underlying this goal included:

Aim #1. Use previous information regarding human and automation performance in this same task and environment as ground truth to facilitate real-time prediction of the relative performance of the agents as a function of the course and scenario characteristics (i.e. perturbation onsets, nearness to ambient traffic, etc.).

Aim #2. Compare the "natural", or un-managed, usage behavior of individuals to their usage decisions when provided visual feedback regarding predicted performance in real-time as an individual case where handoff and takeover behaviors can be influenced in real-time to the end of improving overall performance.

Aim #3. Compare two different control algorithms for providing the predictive feedback to the individual subjects, one that simply reports on relative likelihood of success and one that weights the likelihood according to the potential consequence of failure given the current scenario configuration. The expectation is that the probability of success alone, which should vary more dynamically, will not be as useful to the participants as will the inclusion of consequence, which will modify the degree to which visual feedback indicates a need for change based on current consequence level (e.g. during low consequence instances, the indicator will not as strongly indicate a need for a specific agent to be in control as compared with a high-consequence circumstance where the change is more essential to success). In particular, we expect the user to show higher use of the consequence-based controller and, as a resultant, better performance and ultimately, greater trust in the overall system.

Aim #4. Test a prototype of a real-time predictor based on physiology, behavior, and context variables that has very recently become available to us. As this is a first examination and to avoid contamination of the main experiment, this will only be assessed in a limited subset of the participant population pending date of the delivery of the real-time implementation. Thus an added experimental condition, which will not increase the overall timeline, will be included for this subset of participants.

> Note: A separate pilot study, using the same experimental paradigm and similar apparatus, was completed to address this research aim, but those data are not included with the ARL_ARPI_TX22 dataset.

2.5.1.2.1 Protocol

The research protocol information for this study is as follows:

> Organization: ARL
> Protocol Number: ARL-16-045
> Protocol Name: "Predicting and Influencing Trust-Based Decision Making during Human Interaction with Driving Automation ('TX22')"
> Contract: W911NF-10-D-0002

2.5.1.2.2 Location

This study was conducted at GVSC's Ground Vehicle Simulation Laboratory (GVSL) in Warren, MI, using the Ride-Motion Simulator (RMS).

2.5.1.2.3 Subjects

A total of 19 subjects participated in this study. Each subject completed all tasks in a single day. A total of 68 data recordings were produced from 17 subjects, including 17 recordings where only manual control of the vehicle was available and 51 where the different levels of autonomous vehicle control were available. No useable data was recorded for two of the 19 subjects.

2.5.1.2.4 Apparatus

The following items were used during the experiments. All equipment is located within the GVSL at GVSC.

Ride Motion Simulator (RMS): The RMS is a 6 degree-of freedom (DOF) motion-based simulator that is capable of reproducing the dynamics of military ground vehicles over a vast array of terrains that are commonly experienced by current force vehicles. As shown in Figure 2-48, the RMS is composed of a platform that is mounted on a hexapod design of hydraulic actuators, the coordinated motion of which produces precise resultant motions through 3 cardinal linear (longitudinal, lateral, and vertical) and 3 rotational (yaw, pitch, roll) dimensions. The RMS supports a re-configurable cab that allows the experimenter to customize the driver controls, incorporate rugged display monitors, seats, and data collection devices (e.g. eye tracking and EEG). Supported by the Real-Time Simulation Framework, this environment allows the user to be immersed into a synthetic terrain of various types and experience realistic ground vehicle dynamics. It has integrated motion, audio, and visual systems for high fidelity real-time simulation.

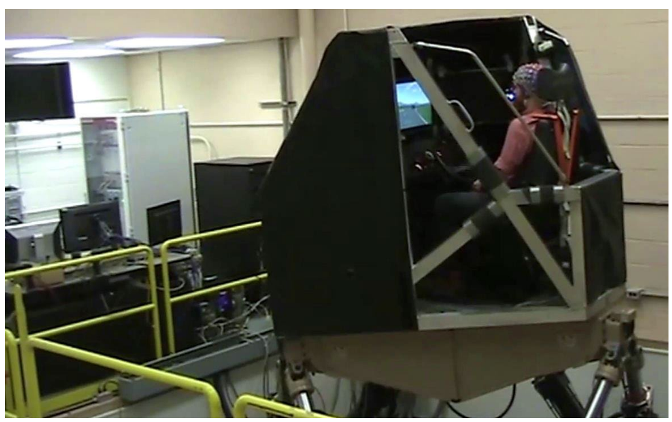


Figure 2-49. Ride Motion Simulator

Crewstation Cab: The simulation environment was constructed to present participants with visual, motion and audio cuing in order to recreate a realistic driving experience. For this experiment, the cab was configured as a single-occupant crew station with 3 flat panel displays, thus allowing creation of central and peripheral visual flow patterns. The simulator cab was comprised of a vehicle seat with a 4-point harness, a steering buck with wheel and standard gauges (speedometer, tachometer, etc.), throttle and brake controls on the floor, and custom controllers for additional task inputs. Audio cuing was limited to presenting the participant with the experimenter's voice (during instruction and other communication periods), the vehicle's sounds (engine noise correlated to engine RPM), and driving-related environmental sounds (e.g. wind gusts, passing vehicles).

Electrophysiological Data Recording: Electrophysiological data was recorded using a commercially-available Biosemi ActiveTwo system. Electroencephalographic (EEG) data was collected using an electrode cap with 64 active surface electrodes. The ActiveTwo system uses sintered Ag/AgCl electrodes held in place by a snug elastic fabric cap. This system does not require conventional skin cleaning and preparation, such as abrasion, but does require gel; hypoallergenic, bacteriostatic gel was used to establish conductive contact. The procedure for placing the electrodes was standard, minimally discomforting, and has been extensively used in both clinical practice and research. The configuration of the electrophysiological recording system included 5 bipolar pairs: vertical (VEOG) and horizontal (HEOG) electrooculogram, electrocardiogram (ECG), 2 natively referenced channels (left and right mastoid process), and one EDA channel. The VEOG pair was centered superior and inferior to the left eye. The HEOG pair was placed along the outer canthus of each eye. The ECG pair was placed on the notch atop the sternum and clavicle of the participant. The bipolar EDA was placed on the palmar surface of the distal phalanges or the thenar and hypothenar eminences of the left or right hand (whichever hand is not the one that the participant identifies as her/his primary steering hand).

Head and Eye Tracking: Minimally calibrated eye gaze and head movement data was acquired using a commercially-available eye tracking system (Smart Eye AB) comprising infrared emitters and four computer controlled optical zoom cameras that allow 6 degree-of-freedom head tracking, 2 degree-of-freedom gaze tracking, and eye blink detection and pupil diameter monitoring.

Video/Audio Recording System: Video data was captured using 2 video cameras mounted on the inside of RMS cab, one aimed at the driving display and the other at the subject's face. All video data was used for real-time subject monitoring and is available to facilitate interpretation of quantitative data. Audio data was recorded in order to capture dialogue between the participant and investigators.

Standard PCs, Monitors, and Input Devices: This project required connecting an array of standard PCs as a local network to enable data synchronization and storage. The RMS control and simulation framework used a small cluster of high-speed computational and network assets, with software that was optimized for real-time operations with low latency, yet able to preserve physics-based fidelity of vehicle and terrain models. Additional PCs were used for audio/video recording, eye tracking, surface electrophysiology (EEG/EOG/ECG/EDA), and for the presentation of task stimuli. Responses to pedestrian appearances were made using a commercially-available game controller (the Nintendo® Wii™ remote).

2.5.1.2.5 Demographics

Demographic information collected for this dataset includes subject ID, age, and dominant hand, however, this data was not processed in the current version of the data. It's available in the raw data only.

2.5.1.2.6 ARL_ARPI_TX22 Tasks

Subjects performed a set of tasks that represented challenges typically associated with driving. In the majority of the experiment conditions, many aspects of the driving tasks were performed jointly with a driving automation. Task objectives included maintaining lane position, maintaining a prescribed following distance behind a lead vehicle, and monitoring for and responding to pedestrians that appear on the side of the road. Each of these objectives had specific associated challenges, to be described below. The pedestrian monitoring task provided a persistent, high-consequence demand throughout each condition that encouraged the subjects to use the automation as much as possible, thus pressuring subjects to continuously make active decisions whether to trust the autonomy.

For each condition, subjects were instructed to drive a simulated vehicle (denoted "ownship") 1.5 laps around the two-lane course while maintaining a distance within a prescribed range (10 – 30 m) behind the lead vehicle. Figure 2-49 illustrates the complete driving course.

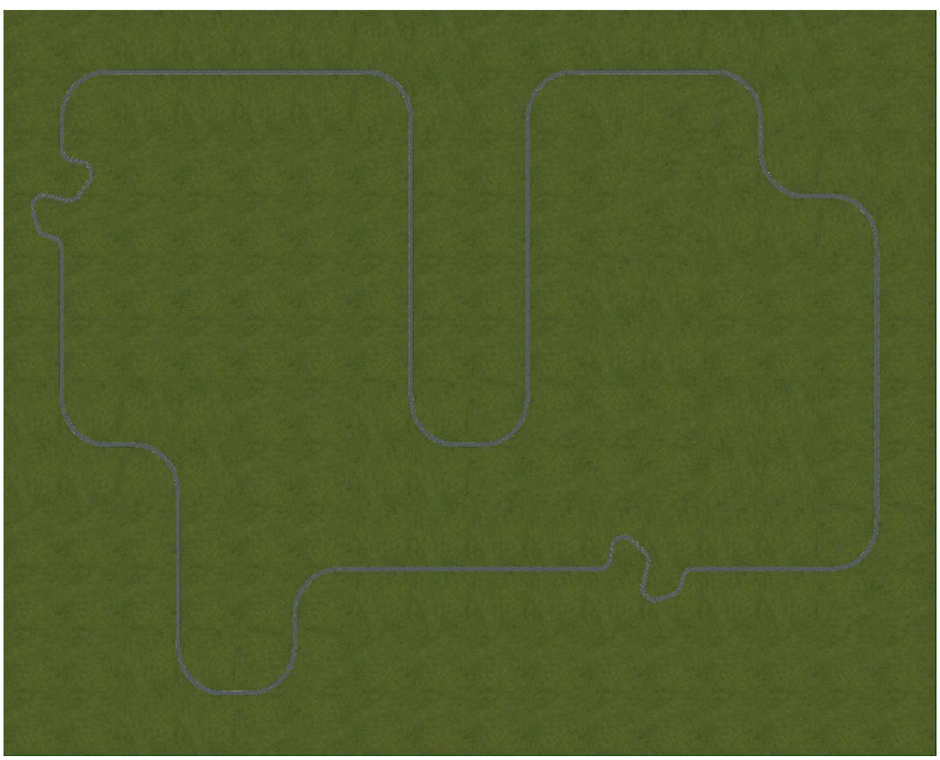


Figure 2-50. Driving Course for ARL_ARPI_TX22 Experiments

Participants were given screen feedback indicating current settings, error occurrences, and score. The screen display presented to the participant is shown in Figure 2-50. Speed was determined by the lead vehicle and road characteristics (i.e. curves). By design, the lead vehicle maintained an average speed of 25 mph, with perturbations and curves leading to speeds temporarily varying as low as 10 and as high as 30 mph. Moreover, participants were instructed that they must not deviate outside of their lane and that all collisions should be avoided. While the driving automation (here forward called "the automation") was capable of controlling vehicle speed and steering, it was not capable of avoiding collisions. In conditions where the automation was available, subjects always had the option to enable it using a toggle switch (enable) with their left foot or disable it using either the throttle or brake (disable). An indicator on the screen provided constant status of the automation, a large arrow pointing to one of two mode indicators; with a large circled "A" for "auto" and "M" for manual. When the automation was enabled, the subjects were instructed to allow the automation to control the throttle, brake, and steering wheel. In addition to automation conditions, there was a manual driving baseline with no available automation, thus the subjects had to complete the entire task themselves. Note the dot between the "A" and "M' indicators in Figure 2-50. This dot moved between the two indicators and indicated the probability of success for an automation selection. If the dot was far to the left, it indicated a high probability of success if automation was selected, if far to the right, it indicated a higher probability of success if the subject retained manual control.

Figure 2-51. ARL_ARPI_TX22 LFD Screen Presentation

Figure 2-51 provides a summary of the LFD paradigm. Subjects drive a vehicle ('ownship', green) while following a lead vehicle (orange). Subjects are instructed to maintain a particular following distance (horizontal dotted line) as well as consistent lane position (vertical dotted line). In some conditions, the subject shares vehicle control with an auto-driver of a performance reliability defined as represented by the distributions labeled σ_L and σ_F. Environmental challenges are presented in the form of ambient traffic (f_T), perturbations to ownship heading (P_H) and lead vehicle speed (P_S), and static and dynamic pedestrians which appear with a specified frequency (f_p) and probability structure (probability of dynamic pedestrian, ρ_d).

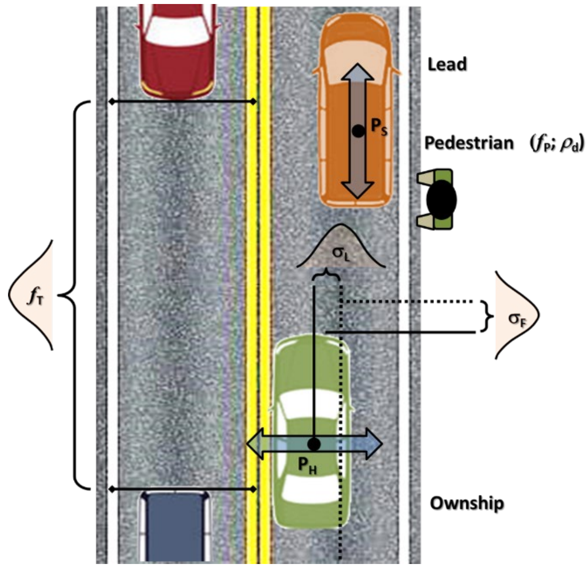


Figure 2-52. Summary of the LFD paradigm for ARL_ARPI_TX22

The autonomy had good, but imperfect, performance reliability defined by distributions shown and labeled as σ_F and σ_L in Figure 2-51. These distributions define the amount of variability in the automation's maintenance of following distance and lane position, respectively. In Figure 2-51 the desired average lane position / following distance is shown with the dotted line and the current position of the ownship is indicated with the solid line; the distributions represent the variance of the deviations around the dotted lines (the difference between the dotted and solid lines). For each performance variable the narrow distribution represents the variance allowed by the automation. Moreover, during the driving task, subjects experienced vehicle perturbations in both lateral and longitudinal directions. Lateral, or heading, perturbations (P_H) were presented as gust-like forces acting to push the vehicle out of its lane. Longitudinal, or speed, perturbations (P_S) involved changes in the speed of the lead vehicle through sudden positive (throttle) or negative (brake) accelerations changing speed by 5 miles per hour in either direction (other speed changes, such as at tight s-curves, were larger than those due to these pseudo-random perturbations). These perturbations occurred, on average, once every 15 ± 5 sec (normally distributed), resulting in approximately 4-5 perturbations per minute and approximately 72 - 90 per 18 minute condition; no perturbations overlapped in time, meaning that a lateral and longitudinal perturbation did not occur simultaneously. The perturbations were semi-randomly selected as to type (lateral/longitudinal) and direction (positive or negative along relevant axis). During the task, any collisions with other vehicles or pedestrians did *not* have a motion consequence (the RMS shutter as if driving over a "rumble strip" on the road), a visual collision indicator appeared on the screen.

Our prior experience with simulated driving has shown that participants have difficulty resolving and responding to multiple sequential perturbations in a longitudinal direction; this is because it can be difficult to distinguish the end of one speed perturbation and the start of another.

Therefore, there were at least two lateral perturbations for each longitudinal perturbation. In order to encourage as much use of the automation as possible, subjects also performed a task for which there were no responses from the automation. This task was designed to be relevant and similar to typical driving in that it involved reacting to pedestrian entities, some of which stepped directly into the pathway of the vehicle. Although this task was unlike real-world driving in that the pedestrians "popped" into view in the scenario, it did reflect more infrequently occurring natural events such as a bicyclist or runner that suddenly emerges from behind a parked car. The pedestrian task was chosen to be presented in this fashion in order to facilitate analyses that have experimental analogs, such as the classic "oddball" task (see Garrido et al., 2009 for review and Squires et al., 1975 for original publication), within standard laboratories that study cognitive function. Throughout all driving conditions, pedestrians dynamically appeared on either the left or right shoulder of the road with a 50% probability; 15% of the pedestrians (the rare or "oddball stimulus") immediately began walking into the path of the vehicle when they appeared. Subjects were asked to respond to *all* pedestrians using a button press on a commercially-available game controller (the Nintendo® Wii™ remote), with static and dynamic pedestrians assigned to the thumb and forefinger of the same hand, respectively. This was kept the same across all participants because, based on pilot testing, it was decided that the differential control between thumb and forefinger was enough to preserve across subjects rather than counterbalance and potentially wash out effects. The response device was held in the same, subject-selected hand throughout all conditions of the experiment and the other was used for control of the steering wheel. Pedestrians appeared at approximately double the rate of vehicle perturbations; specifically, a new pedestrian appeared every 6 ± 1 sec. Once they appeared, a pedestrian remained in the scenario until the vehicle drives past it (approximately 1.5 sec) or until the participant responds with a correct button press, in which case the pedestrian simply disappeared. It was possible to "collide" with a pedestrian. However, these collisions were not represented with any physical consequences; the vehicle simply drove through the pedestrian image and an error icon was displayed on the screen. Moreover, if the initial button press response of the subject was incorrect, they had the opportunity to clear the pedestrian out of the way with a successive correct button press. Therefore, an incorrect response to a pedestrian did not necessarily mean that the ownship would collide with that pedestrian. Figure 2-52 shows the display following a pedestrian collision. The red error icon third from left indicates a pedestrian collision has occurred. The illuminated icon fifth from left indicates that an incorrect button push also occurred.

Figure 2-53. Display Following a Collision with a Pedestrian for ARL_ARPI_TX22

Including the two error icons discussed above, a total of eight circular icons were visible on the presented display, as shown in detail in Figure 2-53. These icons were mostly transparent until a relevant event occurred, at which time they illuminated.

Figure 2-54. ARL_ARPI_TX22 Display Icons

From left to right, the icons indicated the following:

- High range consequence zone. See below.
- High lane consequence zone. See below.
- Collision with a pedestrian.
- Collision with another vehicle (oncoming or lead vehicle).
- Incorrect pedestrian identification button push.
- Lane departure/violation. Yellow if out of lane limits, flashes red if a violation (points loss) occurs for remaining out of lane in excess of 2 seconds.
- Range departure/violation. Yellow if out of range limits, flashes red if a violation (points loss) occurs for remaining out of range in excess of 2 seconds.
- Zone completion. Indicates bonus accrued for completing a zone.

Performance was quantified according to the following scoring structure. Participants began each LFD task with 500 pts. For the most part, points were not gained, but only lost as a consequence of performance errors. Score feedback was provided in real time using minimally-disruptive, centrally-presented icons that illuminated when an error occurred, as well as a persistent score bar across the bottom center of the central display. Each collision with a pedestrian, categorization error through incorrect button press responses, collision with other vehicles, and driving errors leading to lane departures or violations of the following distance threshold, incurred penalties dependent on the consequence zone of the course; in the nominal condition, these losses were -100, -5, -50, and -2 per 2 seconds in error, respectively.

Each condition involved pre-selected course segments that vary in consequence, which means that point values for lane and range violations changed at specifically chosen places in the course. For a high lane consequences zone, point penalties for range were halved (-1 per 2 seconds) while point values for lane violations were doubled (-4 points per 2 seconds). The converse scoring change occurred in high range consequence zones; meaning range violations were double while lane violations were halved. It was always the case that a nominal consequence zone occurred before successive lane or range consequence zones. The intention of these zone changes was to create specific circumstances wherein either agent (human or automated) is preferred; for instance, humans are generally superior at handling lateral perturbations and automation is superior at maintaining following distance in light of sudden speed changes. In all cases, visual feedback was provided to indicate to the subject the current consequence level using on-screen icons (Figure 2-53, far left). In the nominal case, both consequence indicators remained transparent. There were 4 sequences of consequence-course mappings. All subjects had the same consequence sequence for the manual (no automation) condition, and the remaining sequences were counterbalanced across the experimental conditions. This provided for a degree of unpredictability in consequence changes by the subjects.

Because of the nature of the lane and range violations as compared with the others, meaning they are not entirely discrete and may be corrected for a short period of time prior to the point deduction (i.e. it takes 2 seconds of the violation to persist before a deduction occurs), the indicator behavior was also nuanced. That is, the indicators for lane and range provided a warning (yellow color) when the condition existed for point deduction and then briefly flashed red when a point loss occurred. Figure 6 demonstrates the behavior of the lane and range indicators. The top row of icons provide a warning when the conditions exist for a potential deduction, but the time requirement (2 seconds) has not been passed. When a point loss has occurred, the icon briefly (~0.7 sec) changed from yellow to red to signal the loss and then returned to yellow or, if the condition was corrected, neutral.

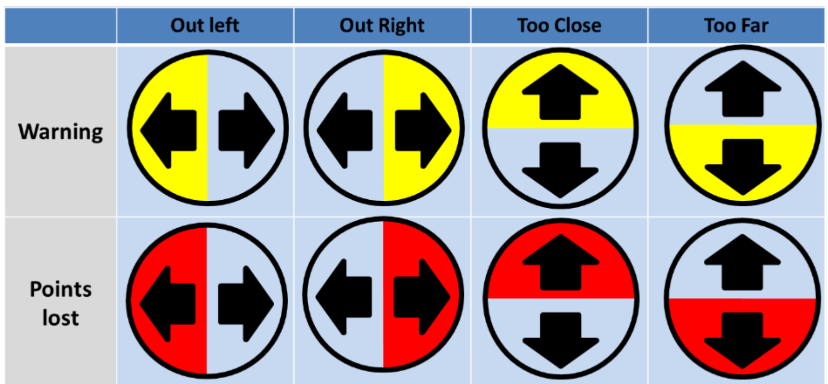


Figure 2-55. Lane and Range Indicators

To discourage subjects from completely giving up on the task in light of a poor performance causing loss of all points, a small bonus (5 points regardless of consequence) was given at the completion of each of the zones throughout the driving course. As there are 48 total zones in the course, this bonus could offset point losses of up to 240 points per condition. The final (total) score for the experiment determined an amount of "bonus" pay received by the participants. All participants were guaranteed at least a 50% bonus on their performance, however they were be made aware of this in advance.

2.5.1.2.7 POC

Jason Metcalfe (jason.s.metcalfe2.civ@army.mil), ARL, Primary Investigator
Tony Johnson (tjohnson@dcscorp.com), DCS Corp., Data Engineer
Michael Dunkel (mdunkel@dcscorp.com), DCS Corp., Data Engineer

2.5.2 ARPI Publications

Drnec, K., Gremillion, G., Donavanik, D., Canady, J.D., Atwater, C., Carter, E., Haynes, B.A., Marathe, A.R., & Metcalfe, J.S. (2018). **The role of psychophysiological measures as implicit communication within mixed-initiative teams**. In J.Chen and G. Fragomeni (Eds.). *Virtual, Augmented, and Mixed Reality: Interaction, Navigation, Visualization, Embodiment, and Simulation. VAMR 2018. Lecture Notes in Computer Science, 10909*.

DOI: *10.1007/978-3-319- 91581-4_22*.

Drnec, K., Marathe, A.R., Lukos, J.R., & Metcalfe, J.S. (2016) **From trust in automation to decision neuroscience: Applying cognitive neuroscience methods to understand and improve interaction decisions involved in human automation interaction.** *Frontiers in Human Neuroscience, 10*, 290.

DOI: *10.3389/fnhum.2016.00290*.

Drnec, K. & Metcalfe, J.S. (2016). **Paradigm development for identifying and validating indicators of trust in automation in the operational environment of human automation integration.** In D.D. Schmorrow and C.M. Fidopiastis (Eds). *Foundations of Augmented Cognition 2016, Lecture Notes in Computer Science, 9744*, 157-167.

DOI: *10.1007/978-3-319-39952-2_16*.

Gremillion, G.M., Metcalfe, J.S., Marathe, A.R., Paul, V.J., Christensen, J., Drnec, K., Haynes, B., & Atwater, C. (2016). **Analysis of trust in automation for convoy operations.** *Proc. SPIE, 9836*, Micro- and Nanotechnology Sensors, Systems, and Applications VIII, 98361Z.

DOI: *10.1117/12.2224009*.

Metcalfe, J.S., Marathe, A.R., Haynes, B., Paul, V.J., Gremillion, G.M., Drnec, K., Atwater, C., Estepp, J.R., Lukos, J.R., Carter, E.C., & Nothwang, W.D. (2017). **Building a framework to manage trust in automation.** *Proceedings of the SPIE, 10194*, Micro- and Nanotechnology Sensors, Systems, and Applications IX, 10194U.

DOI: *10.1117/12.2264245*.

Neubauer C., Gremillion G., Perelman B.S., La Fleur C., Metcalfe J.S., Schaefer K.E. (2020) **Analysis of facial expressions explain affective state and trust-based decisions during interaction with autonomy.** In: Ahram T., Karwowski W., Vergnano A., Leali F., Taiar R. (eds) *Intelligent Human Systems Integration 2020. IHSI 2020. Advancesin Intelligent Systems and Computing, vol 1131*. Springer, Cham.

DOI: *10.1007/978-3-030-39512-4_152*

Nothwang, W., Gremillion, G.M., Donavanik, D., Haynes, B.A., Atwater, C.S., Canady, J.D., Metcalfe, J.,& Marathe, A. (2016). **Multi-sensor fusion architecture for human-autonomy teaming.** *2016 Resilience Week.*

DOI: *10.1109/RWEEK.2016.7573327*.

Orellana, C. B., Rodriguez, L. R., Gremillion, G. M., Huang, L., Demir, M., Cooke, N., Metcalfe, J. S., Amazeen, P. G., & Kang, Y. (2022). **The impact of automation conditions on reliance dynamics and decision-making.** *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 66(1), 721–725.

DOI: *10.1177/1071181322661477*.

Rodriguez Rodriguez, L., Bustamante Orellana, C., Gremillion, G.M., Huang, L., Demir, M., Cooke, N., Metcalfe, J.S., Amazeen, P.G., & Kang, Y. (2022). **Performance and Relative Risk Dynamics during Driving Simulation Tasks under Distinct Automation Conditions.** Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 66(1), 1230–1234.

DOI: *10.1177/1071181322661471*.

Rodriguez Rodriguez, L., Bustamante Orellana, C., Landfair, J., Magaldino, C., Demir, M., Amazeen, P. G., Metcalfe, J. S., Huang, L., & Kang, Y. (2021). **Dynamics of trust in automation and interactive decision making during driving simulation tasks.** *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 65(1)*, 786–790. DOI: *10.1177/1071181321651288*.

2.5.3 ARPI Datasets

A total of 187 datasets were collected across 41 recording sessions.

The total size of the EEG files is 132.21 GB, representing 46.02 hours of EEG recording.

2.5.3.1 ARL_ARPI_TX20 Dataset

A total of 119 datasets were collected across 24 recording sessions, from 24 subjects recruited from the Warren, Michigan area. The 119 datasets include 16 of the PRACTB task, 23 where only manual control of the vehicle was available, and 80 where some level of autonomous control was available. A complete set of vehicle control tasks (one manual control and four with some level of autonomous control) was collected for 17 subjects.

The total size of the EEG files is 85.26 GB, representing 26.32 hours of EEG recording.

2.5.3.2 ARL_ARPI_TX22 Dataset

A total of 68 datasets were collected across 17 recording sessions, from 17 subjects recruited from the Warren, Michigan area. A complete set of recordings (one task where only manual control of the vehicle was available and three tasks with some level of autonomous control was available) was collected for all 17 subjects.

The total size of the EEG files is 46.95 GB, representing 19.70 hours of EEG recording.