Department of Transportation
Federal Railroad Administration

1**INFORMATION COLLECTION SUPPORTING STATEMENT B**
Design and Evaluation of a Robust Manual Locomotive Operating Mode
OMB CONTROL NUMBER 2130-NEW

1. Description of sampling method to be used.

Human subjects will be selected on the basis of locomotive operating experience representing the "experienced" subject group and college students will be selected from systems engineering and other classes to represent the "novice" subject group. Expert subjects will be recruited through GE's contacts with Class 1 railroad senior management. These subjects will be expected to have experience with standard automation equipment in broad use in the industry. Novice participants will be recruited through an announcement given in relevant classes (e.g., systems engineering, human factors, etc.) and posted on bulletin boards (e.g., near MVL offices). Participant information which are theoretically relevant to their task performance (e.g., year experience in railroad operations, experience with video games, CPRS, etc.) will be collected from participants to account for any confounding factors. The recruitment plan and study parameters will be reviewed by MIT Institutional Review Board (IRB) in detail prior to execution.

2. Description of procedures for information collection, including statistical methodology for stratification and sample selection.

Each participant will be required for approximately 9.0 hours. The following protocol for the study will include 10 college students and 20 freight locomotive engineers. The estimated sample size is based on similar studies in the research literature and the research team's previous experience with human-automation studies. Dependent variables will measure crew performance (e.g., include safety measures such number of speed violations, duration of head-down time, probability of detection and time to detection, as well as system efficiency (e.g., fuel burn, time to arrival). Metrics will also be evaluated that reflect the effectiveness of the collaborative control system under development, including situation awareness (e.g., using the situation awareness global assessment technique and operator workload (e.g., using subjective ratings and/or objective secondary task performance. In addition, control variables will include individual differences (e.g., as measured by complacency potential rating scale, boredom proneness, and cognitive failure. Statistical test (e.g., regression/ANOVA) will be performed on the data to examine the hypothesis that the new operating mode will reduce the performance gap between expert and apprentice participants

The operational scenario used in the experiment sessions described below will be representative of typical freight operations and will be derived from the research team's previous study and consultation with subject matter experts. The prototyped display is intended to be designed to assist novice locomotive drivers, represented by the college students. There are many new inexperienced drivers entering service as many more experienced ones retire. The college students will represent novice or apprentice locomotive engineers with little or no experience. Their performance will not be directly compared to the experienced engineers in the study. Rather, their performance from a usability standpoint with the prototype display is what is of interest in the study. Their performance will be used as an indicator of how well a new engineer might perform with the prototyped display and they will be used to pilot the data collection methodology. Similarly, the performance of experienced engineers and their usability of the prototype is of interest.

The first part of the experiment session will be a training session lasting approximately three hours for each participant, so 90 hours in all. Subjects will receive a brief introduction to the Cab Technology Integration Lab (CTIL), including how to operate the simulator and a description of the train route and nominal train schedule that they will drive during the experiment session.

After the training, subjects will begin an experiment session that is expected to take about six hours each, so 180 hours in all. Subjects will drive the test route a total of four times within a single day. This estimate includes breaks and lunch. This is identical in duration to a study performed by the research team in 2015-16. The study parameters and design will be reviewed by MIT COUHES IRB in detail prior to execution (as was the prior study).

The order of the conditions will be balanced across the subjects, although the project may be constrained by the total number of subjects that are recruited for the study. Drivers will be given an itinerary to follow for each of the four driving sessions which will include many typical events that occur en route such as meets and passes, scheduled stops at specific locations (i.e., to change the consist), or speed restriction changes due to maintenance of way. The timing and location along the route of these events will differ between the conditions to ensure that subjects do not learn the scenarios and anticipate the critical moments in the test.

During each test session, the subject will also receive instructions from the experimenter (acting as a dispatcher) outlining changes in operating condition. Drivers are expected to remember these changes to act at the appropriate time further down the track. In the test conditions using automation, subjects will have to interact with the automation as needed to ensure the correct control behavior is applied.

At the times when the engineer is waiting for an event to clear or pass, the subject will be asked questions to test their current situation awareness and ability to project to future conditions, such as, “What is your current location?” and “What is the distance to the next stop?”

Workload will also be assessed during this pause by the subject rating his/her workload by completing the NASA-TLX subjective rating form to confirm workload measurements derived from the scenario simulated operating task. The rating form will take approximately 1 minute to complete times 30 subjects x four (4) 4 scenarios or completions, yielding 120 minutes, or about 2 hours in all.

Subjects' driving performance will be measured by the adherence to operational rules, adherence to the schedule, and a measure of situation awareness. After completing the experiment, subjects can ask any further questions about the experimental technology but will be free to leave.

3. Description of methods to maximize response rate and to deal with non-response issues.

Each Class 1 railroad will be asked to provide a small number of participants; they have historically been willing to participate. The performance measures taken during the study are all taken from the research literature and are similar in nature to the research team's previous study.

4. Describe any test procedures for procedures or methods to be undertaken.

The test materials and procedures will build on those used for the research team's prior study.Namely, GE has been involved in the development of automation technologies at various levels of rail operations for many years. Systems include slow speed control (for coal loading), remote control operations, and Trip Optimizer. In addition, GE and MIT have recently developed a functional task model, and a methodology for using it, to inform the design of automation systems16. The MIT Man Vehicle Laboratory has conducted several projects that have investigated the effects of new technology on operator performance, including FAA sponsored projects on predictor displays, evaluating new alerter technology and novel moving map and HUD displays, and NASA/NSBRI sponsored research on human-automation task allocation, the effects of changing automation modes on operator performance, workload and situation awareness, and automated checklist design. Several projects utilized Hierarchical Task Analysis, MATLAB/Simulink and/or object process methodology for quantitative modeling.

Limitations

This is primarily an exploratory study of the feasibility of an innovative cab technology that may narrow the gap in performance between novice and experienced locomotive engineers. Given this, the work includes several limitations that may affect the generalizability of the results. First, this is a convenience sample utilizing thirty (30) locomotive engineers which is not a sample size that would be large enough to represent the population of all locomotive engineers in US operations. Also, this study will be conducted in a simulator and hence the results cannot necessarily be directly generalized to the railroad operating community without further study. Future studies would need to be conducted to further investigate this innovation.
The FRA will not be making recommendations as a result of this study and the FRA will be publishing the study independent of whatever results are obtained.

5. Provide name and phone number of individuals consulted on statistical aspects of study design and other persons who will collect/analyze information for agency.

No additional consultation beyond original research proposers. However, as required by 45 CFR 46, *Protection of Human Subjects*, the overall study will be presented to an Institutional/Independent Review Board for approval of the study design, collection of experimental data, and procedures for the protection of human subjects.

The total estimated cost to the Government to complete the 2-year study is $990,000 under a cost-share agreement with the researchers to design and develop the prototype system in the first year and to test and evaluate (with human subjects) in the second year.

The point of contact for the study will be:

Michael E. Jones
US Department of Transportation
Federal Railroad Administration
Office of Research and Development
Human Factors Division
Washington, DC. 20590
Michael.e.jones@dot.gov
202-493-6106