DEPARTMENT OF TRANSPORTATION

INFORMATION COLLECTION SUPPORTING STATEMENT: PART B

TITLE OF INFORMATION COLLECTION: Human Interaction with Driving Automation Systems

OMB CONTROL NUMBER: 2127-NEW

ABSTRACT[1]

The National Highway Traffic Safety Administration (NHTSA) has proposed to perform research involving the collection of information from the public as part of a multi-year effort to learn about how humans interact with driving automation systems (DAS). This research will support NHTSA in understanding the potential safety challenges associated with human-DAS interactions, particularly in the context of mixed traffic interactions where some vehicles have DAS and others do not. Within mixed traffic environments, vehicles may also have DAS that perform more or less of the driving task (i.e., different levels of automation) and come with their own sets of expectations and limitations. This research will add to the state of knowledge and is not immediately intended to inform regulations or policy.

The data collections will be performed once to obtain the target number of valid test participants. Study participants will be members of the general public and participation will be voluntary with monetary compensation provided. Participants are generally healthy individuals aged 18 and older, recruited using the DSRI registry and through email blasts to University of Iowa community. Efforts will be made to enroll a diverse age sample that broadly represents the age of the U.S. driving population and includes those at greater risk of crashing (e.g., less than 25 years of age and greater than 65 years of age).

The research will be conducted in three parts, referred to as Study 1, Study 2, and Study 3. All study procedures will be approved by the University of Iowa Institutional Review Board (IRB). Data collection will begin upon receipt of PRA clearance and will involve human-subjects data collection using the driving simulators at the University of Iowa Driving Safety Research Institute (DSRI).

The objective of the first study is to understand how humans interact with DAS in mixed traffic environments. In the first study, participants will participate in pairs with each participant driving a separate driving simulator but interacting in the same driving environment. Participants will experience one of two driving automation systems. Both members of the participant pair will provide informed consent, a pre-drive questionnaire, a training presentation, a familiarization drive, wellness questionnaires to screen for simulator sickness, a study drive, in-drive ratings of trust, and a post-drive questionnaire and risk-propensity assessment. During the simulator drives, one member of the pair will perform a continuous drive along a specified route. The other member of the pair will complete three short drives where they interact with the other participant at specific points throughout the drive. The simulator will collect vehicle data (e.g., brake inputs, steering wheel angle) and data about the surrounding environment (e.g., distance to surrounding vehicles and lane markings). After the drives, participants will complete a questionnaire to assess their understanding of the DAS and their trust in and acceptance of the DAS. Data will be analyzed to understand how human drivers interact with DAS in mixed traffic situations and to understand how humans understand and perceive automation in different situations.

---

[1] The Abstract must include the following information: (1) whether responding to the collection is mandatory, voluntary, or required to obtain or retain a benefit; (2) a description of the entities who must respond; (3) whether the collection is reporting (indicate if a survey), recordkeeping, and/or disclosure; (4) the frequency of the collection (e.g., bi-annual, annual, monthly, weekly, as needed); (5) a description of the information that would be reported, maintained in records, or disclosed; (6) a description of who would receive the information; (7) the purpose of the collection; and (8) if a revision, a description of the revision and the change in burden.

In the second study, participants will complete a drive in a driving simulator with a driving automation system. The study drive will contain situations to which the DAS must respond. Participants will be randomly assigned to one of three systems with different capability, defined by how well the automation can navigate the set of test situations. The simulator will collect vehicle data (e.g., brake inputs, steering wheel angle) and data about the surrounding environment (e.g., distance to surrounding vehicles and lane markings). After the drives, participants will complete a questionnaire to assess their understanding of the DAS and their trust in and acceptance of the DAS as well as a risk-propensity assessment. Data will be analyzed to understand how human drivers interact with DAS in mixed traffic situations and to understand how humans understand and perceive automation in different situations.

In the third study, participants will complete a drive in a driving simulator with a driving automation system. The study drive will contain situations to which the DAS must respond. Participants will be randomly assigned to one of three systems with different capability, defined by how well the automation can navigate the set of test situations. Outside of this, study procedures are the same as those for the second study.

Part B. Statistical Methods

1. <u>Describe potential respondent universe and any sampling selection or other respondent selection method to be used</u>.

Respondents will be licensed drivers, aged 18+, in Eastern Iowa willing to drive to the University of Iowa Research Park to participate in a driving simulation study. The University of Iowa Driving Safety Research Institute (DSRI) currently has a registry of approximately 7,000 individuals who have already expressed interest in participating in driving research studies. From the registry, potentially eligible individuals (those who meet initial age and sex criteria via registry query) will be contacted via email. If insufficient interest exists from that group, advertisements will be posted to increase the pool of potential participants. Individuals will be selected generally based on order of response and ability to meet study timeline. Efforts will be made to enroll a diverse age sample and a balance of males and females, while being inclusive of non-binary individuals.

No statistical methods will be used in selecting study participants. Study participants will be selected based on a set of criteria that serve to ensure that participants will be generally representative of average U.S. drivers. The criteria state that participants must

a. Be able to attend one visit that lasts approximately 2.5 hours;
b. Be comfortable with and willing to engage in an email task while driving the simulator;
c. Agree to abstain from alcohol and use of recreational drugs in the 24 hours prior to the appointment;
d. Agree to be well-rested for the appointment;
e. Possess a valid U.S. driver’s license and have been licensed for at least two years;
f. Drive at least 2,000 miles per year or drive at least once per week;
g. Be able to drive without any special equipment to help them drive, e.g., pedal extensions, hand brake or throttle, spinner wheel knobs, or other non-standard equipment;
h. Hold a valid driver's license without restrictions related to road type (e.g., no interstate or freeway driving), maximum speed (e.g., maximum speed of 35 mph), or licensure (e.g., intermediate license);

i. Be aged 18+;

j. Be in good general health and meet DSRI health screening criteria put in place for the safety of participants and the research team (excludes many conditions, e.g., pregnancy, seizures/epilepsy with episode in past year, brain damage with lingering symptoms or stroke in past 6 months, vestibular concerns, issues with hearing, unmanaged mental health concerns, current or chronic neck or back injury, mobility concerns that would make using the simulator emergency exit a challenge, great propensity for discomfort when using visual display devices).

2. Describe procedures for collecting information including statistical methodology for stratification and sample selection, estimation procedures, degree of accuracy needed, and less than annual periodic data cycles.

No such statistical methods (e.g., stratification) will be employed.

We have determined that 224 total participants across the three studies are necessary to obtain sufficient power. Statistical power refers to the ability to detect a true effect if one actually exists. Anticipating a typical rate attrition, the total number of participants that will likely need to be run to obtain the valid data sets is 300 individuals. This is 180 participants for Study 1 and 60 participants each for Study 2 and Study 3. Given a loss of one subject results in the loss of the participant pair, approximately 40% was added for Study 1; approximately 25% was added for Study 2 and Study 3. To provide for issues that may arise (e.g., connectivity loss between simulators, answering additional questions, participant simulator sickness), we have estimated burdens for a total of 300 participants.

Individuals who previously have expressed an interest in participating in driving studies with DSRI will be contacted by email. Potential participants will complete an online pre-screening questionnaire to determine their eligibility to participate in each study. Each respondent deemed eligible and who agrees to participate will receive a confirmation email after scheduling. The email will include the date and time of their study visit as well as logistical information such as directions and parking. It will also include a contact email address and phone number in case of cancellation or questions and remind participants that they must refrain from alcohol and recreational drug use (e.g., cannabis) in the 24 hours prior to any study visit. Reminder emails with a health check (ensuring the subject is not feeling ill and therefore likely to cancel or have affected performance at the visit) will be sent approximately 24 hours prior to the appointment.

In the first study, participants will participate in pairs with each participant driving a separate driving simulator but interacting in the same driving environment. Participants will experience one of two driving automation systems. Both members of the participant pair will provide informed consent, a pre-drive questionnaire, a training presentation, a familiarization drive, wellness questionnaires to screen for simulator sickness, a study drive, in-drive ratings of trust, a post-drive questionnaire, and a balloon analogue risk task (BART). During the simulator drives, one member of the pair will perform a continuous drive along a specified route. The other member of the pair will complete three short drives where they interact with the other participant at specific points throughout the drive. The simulator will collect vehicle data (e.g., brake inputs, steering wheel angle) and data about the surrounding environment (e.g., distance to surrounding vehicles and lane markings). After the drives, participants will complete a questionnaire to assess their understanding of the DAS and their trust in and acceptance of the DAS, as well as general risk-taking behaviors while driving. Finally, participants will engage with the Balloon Analogue Risk Task and the visit will be complete. Procedures for Study 2 and Study 3 are the same, but only one participant will be driving. In the second study, participants will complete a drive in a driving simulator with a driving automation system. The study drive will contain situations to which the DAS must respond. Participants will be randomly assigned to one of three systems with different capabilities defined by how well the automation can navigate the set of test situations. The simulator will collect vehicle data (e.g., brake inputs, steering wheel angle) and data about the surrounding environment (e.g., distance to surrounding vehicles and lane markings). In the third study, participants will complete a drive in a driving simulator with a driving automation system. The study drive will contain situations to which the DAS must respond. Participants will be randomly assigned to one of three systems with different capabilities, defined by how well the automation system can navigate the set of test situations.

Data from the studies will be collected from several sources, including data from the simulator, video annotation of driver behavior, and through questionnaires. This section provides a summary of each of those measures, with specific data sources and metrics and their analysis described in the subsequent study sections.

- **Simulator driving data** will be collected from the driving simulator(s) and reduced to produce summary measures of takeover, response, and severity. Dependent measures include takeover time, brake response time, steering response time, and time-to-collision. Takeover time will be defined by the time it takes drivers to deactivate the driving automation system. Response times will be defined by the first instances of braking (>5 lbs brake force) and steering (>10 degrees change in steering wheel angle). Severity of the event will be defined by minimum time-to-collision, with negative values indicating a crash.

- **Trust data** will be collected at three points during the study visits. Pre-drive trust will be assessed via a questionnaire. Situational trust will be collected on a seven-point scale during the drives after each study event. Post-drive trust will be assessed via a post-drive questionnaire. Dependent measures include mean pre-drive trust, mean situational trust, and mean post-drive trust.

- **Understanding data** will be collected at two points during the drive, a pre-drive and post-drive questionnaire. The questions measure understanding of different features of the driving automation systems. The questions also ask about confidence in responses. Dependent measures include mean accuracy and mean confidence in the pre- and post-drive questions.

- **Risk-taking propensity data** will be measured after the study drives. Subjective evaluation of risk-taking will be measured via a questionnaire. Objective risk-taking propensity will be measured via the Balloon Analogue Risk Task (BART). Dependent measures include mean score on the risk-taking self-assessment and mean number of pumps aggregated across trials on the BART.

Electronic study data is collected and recorded onto data storage media and will be password protected and accessed only by the principal investigator, study personnel, or system administrators. Data will be transferred to a permanent data storage area at the end of the project where it will only be available to NHTSA or research team members. Simulator data is captured and initially stored on a mirrored RAID system located within a limited access area of the DSRI facility. This data is behind a hardline firewall. Access to study data is controlled through validated user login, authentication protocols, and access permissions established on a per-study basis. Data backups are maintained as dual copies on physical hard drive devices. One drive is stored within a secured location on-site, and the other is stored off-site under the auspices of The University of Iowa Information Technology Services. All backup drives are inventoried and access to study data requires a request for access and authorization from a designated authority. Questionnaires are designed and data are collected online via REDCap. The information is saved in a password-protected account on their website. Data is downloaded by research staff and saved on a password-protected computer located at DSRI. Access to our account is password-protected and only the team can access the questionnaires to view or export the data.

3. Describe methods to maximize response rates.

Participation in the study is voluntary. Response rate will be maximized by initially contacting only individuals who have previously expressed interest in driving research. Participants will be offered $36/hour as compensation for completing all study procedures. To recruit non-college aged individuals, we have to be competitive with hourly rates found in a college town without payment being coercive, yet high enough to also draw in professionals. This rate of pay is comparable to the current national average hourly rate and is also consistent with our experience recruiting study participants. The amount of compensation offered covers typical costs incurred such as travel. Our facility does not charge for parking and the nearest bus is free fare.

In addition to compensation, several methods will be used to maximize response rates, including

- Completing recruiting and screening electronically for respondent convenience,
- Sending a visit confirmation email including the scheduled date and time, directions, pertinent study information, and a contact email address and phone number in case of cancellation or questions, and
- Sending reminder emails within 24 hours of the scheduled study appointment to boost show rate and help alleviate schedule misunderstandings or forgetfulness.

4. Describe tests of procedures or methods.

The project will include three driving simulator experiments.

Driving performance measures will be similar across all three studies and be used to address the individual research questions for each study (below). These sources of data include

1. **Driving data**
2. **Trust data**
3. **Understanding data**
4. **Risk-taking propensity data**

Study 1

In Study 1, pairs of participants (A and B) will interact by driving in two connected simulators in the same virtual traffic environment. The study will be designed such that the participants interact with one another at three points throughout the drive, referred to as study events. The study will be a between-subjects design with level of automation (partial or conditional) as a between-subjects independent variable. There will be four groups, with different combinations of partial and conditional automation within the pair, as shown in Table 1. Partial automation refers to vehicle automation where the human has to monitor the automation while the automated system is controlling the steering, braking, and acceleration of the vehicle. Partial automation does not perform object and event detection and response (e.g., responding to unexpected objects in the roadway). Conditional automation refers to vehicle automation that does not require human monitoring while the automation is controlling the vehicle. Conditional automation performs object and event detection and response while driving within its operational design domain.

Table 1. Study 1 design

	Participant A	**Participant B**
Group 1 (n=32)	Partial	Partial
Group 2 (n=32)	Partial	Conditional
Group 3 (n=32)	Conditional	Partial
Group 4 (n=32)	Conditional	Conditional

*Note sample size represents individual participants (i.e., n=32 equates to 16 pairs).

Study 1 will address the following research questions:

1. **Takeover Time, Response Time, and Severity**

- **Does takeover time differ based on the combination of driving automation systems?**
- **Does response time differ based on the combination of driving automation systems?**
- **Does the combination of driving automation systems impact the severity of an interaction between two drivers?**

Takeover time, brake and steering response time, and time-to-collision will be calculated for each study event. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

Data will be displayed descriptively, i.e., frequencies (%), mean, SD, and ranges. The predictor variables are those from sample characteristics, i.e., age, sex, gender, race, driving experience, trust in automation, risk-taking propensity, and mental model scores. The dependent variables relate to summary scores of the driving performance such as takeover time, response time, and severity. Variables will be assessed for normality to determine use of parametric versus non-parametric analyses via visual examination (e.g., probability plots, histograms, stem and leaf plots) and statistical tests (e.g., Fisher's skewness and kurtosis, Kolmogrov-Smirnov, or Shapiro-Wilks tests). Data will be disaggregated by sex and age. Inferential statistics capture differences between the level of automation (six combinations of partial/conditional pairs) and within-subject (*pre-drive vs post-drive*). Comparisons will be conducted via ANOVA. We will first use ANOVA, with a two-tailed test of significance ($p < .05$), to detect group differences and the Benjamini Hochberg (1995) procedure to identify the differences between levels of a factor, while controlling for multiple comparisons. Next, we will use linear models to regress the independent variables on the dependent variable(s) and to evaluate the effects of one independent variable on the dependent variable while controlling for other (confounding) variables. We will use log-linear models to regress the significant factors based on trust in automation and mental model scores, while controlling for confounders. Trend lines will be plotted to observe change scores for between subjects and within subject comparisons. We will use R (latest version) for data collation, mutation, and analysis.

2. **Trust in Driving Automation Systems**

- **How much do drivers trust driving automation systems before they drive with them and does this vary as a function of level of automation?**

- **How much trust do drivers have in driving automation systems as a function of how they behave in specific events and does this vary as a function of level of automation?**
- **Does driving with a driving automation system increase or decrease trust and does this vary as a function of level of automation?**

Mean trust ratings will be calculated for the pre-drive questionnaire and post-drive questionnaire. Situational trust ratings will be captured for each study drive event. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

For pre- and post-drive trust, a 2 (Level of Automation (partial vs. conditional)) x 2 (Time (pre vs. post)) mixed analysis of variance (ANOVA) will be performed to assess whether trust varied as a function of level of automation before and after the drive. Significant main effects will be examined via paired t-tests for separate dependent measures.

For situational trust, a 2 (Level of Automation (partial vs. conditional)) x 3 (Event) mixed analysis of variance (ANOVA) will be performed to assess whether trust varied as a function of level of automation across the three study drive events. Significant main effects will be examined via paired t-tests for separate dependent measures.

3. **Understanding of Driving Automation Systems**

- **How well do drivers understand driving automation systems after a short training, before they drive with them, and does this vary as a function of level of automation?**
- **Does driving with a driving automation system increase understanding and does this vary as a function of level of automation?**
- **Is there a relationship between takeover time, response time, and response severity as a function of level of automation?**

Mean understanding scores will be calculated for the pre-drive questionnaire and post-drive questionnaire. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

A 2 (Level of Automation (partial vs. conditional)) x 2 (Time (pre vs. post)) mixed analysis of variance (ANOVA) will be performed to assess whether understanding varied as a function of level of automation and whether there was a difference before and after the drive. Significant main effects will be examined via paired t-tests for separate dependent measures.

Correlations will be calculated between mean pre- and post-drive understanding and driving performance measures (takeover time, response time, severity) to determine the relationship between understanding and performance.

4. **Risk Taking Propensity**

- **Does risk taking propensity, measured by both self-ratings and objective assessment, relate to trust in driving automation systems?**
- **Is there a relationship between risk taking propensity, measured by both self-ratings and objective assessment, and takeover time, response time, and response severity?**
- **Is there a relationship between self-rated risk taking and objective risk-taking propensity?**

Mean self-rated risk-taking propensity scores will be calculated from the post-drive questionnaire. Mean number of pumps will be calculated across trials on the post-drive BART. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

Correlations will be calculated between risk-taking self-ratings and mean pre- and post-drive trust and situational trust to determine the relationship between self-rated risk-taking propensity and trust in driving automation systems.

Correlations will be calculated between mean pumps on the BART and mean pre- and post-drive trust and situational trust to determine the relationship between objective risk-taking propensity and trust in driving automation systems.

Correlations will be calculated between risk-taking self-ratings and mean pumps on the BART to determine the relationship between self-rated risk-taking propensity and objective risk taking.

Study 2

In Study 2, participants will drive the NADS-1 driving simulator with a driving automation system. Participants will be randomly assigned to one of six groups (Table 2), with different combinations of driving automation capability (high, moderate, low) and understanding (strong, weak). Automation capability refers to how the automation is able to handle (successfully navigates vs. fails) four different events in the study drive. Level of understanding refers to how well participants understand the functionality and limitations of the driving automation system and is manipulated via pre-drive training presentations.

Table 2. Study 2 design

	Strong Understanding	**Weak Understanding**
High Capability Automation	Group 1 (n=8)	Group 4 (n=8)
Moderate Capability Automation	Group 2 (n=8)	Group 5 (n=8)
Low Capability Automation	Group 3 (n=8)	Group 6 (n=8)

Study 2 will address the following research questions:

1. **Takeover Time, Response Time, and Severity**

- **Does takeover time differ based on the capability of driving automation systems and level of understanding?**
- **Does response time differ based on the capability of driving automation systems and level of understanding?**

Takeover time and brake and steering response time will be calculated for each study event. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

Data will be displayed descriptively, i.e., frequencies (%), mean, SD, and ranges. The predictor variables are those from sample characteristics, i.e., age, sex, gender, race, driving experience, trust in automation, risk-taking propensity, and mental model scores. The dependent variables relate to summary scores of the driving performance such as takeover time, response time, and severity. Variables will be assessed for normality to determine use of parametric versus non-parametric analyses via visual examination (e.g., probability plots, histograms, stem and leaf plots) and statistical tests (e.g., Fisher's skewness and kurtosis, Kolmogrov-Smirnov, or Shapiro-Wilks tests). Data will be disaggregated by sex and age. Inferential statistics capture differences between subjects understanding (strong vs. weak), capability of automation (high, moderate, and low) and within subject (*pre-drive vs post-drive*). Comparisons will be conducted via ANOVA. We will first use ANOVA, with a two-tailed test of significance (p <.05), to detect group differences and the Benjamini Hochberg (1995) procedure to identify the differences between levels of a factor, while controlling for multiple comparisons. Next, we will use linear models to regress the independent variables on the dependent variable(s) and to evaluate the effects of one independent variable on the dependent variable while controlling for other (confounding) variables. We will use log-linear models to regress the significant factors based on trust in automation and mental model scores, while controlling for confounders. Trend lines will be plotted to observe change scores for between subjects and within subject comparisons. We will use R (latest version) for data collation, mutation, and analysis.

2. Trust in Driving Automation Systems

- **How much do drivers trust driving automation systems before they drive with them and does this vary as a function of capability of automation and level of understanding?**
- **How much trust do drivers have in driving automation systems across events as a function of capability of automation and level of understanding?**
- **Does driving with a driving automation system increase or decrease trust and does this vary as a function of capability of automation and level of understanding?**

Mean trust ratings will be calculated for the pre-drive questionnaire and post-drive questionnaire. Situational trust ratings will be captured for each study drive event. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

For pre- and post-drive trust, a 3 (Capability of Automation) x 2 (Level of Understanding) x 2 (Time (pre vs. post)) mixed ANOVA will be performed to assess whether trust varied as a function of capability of automation before and after the drive. Significant main effects will be examined via paired t-tests for separate dependent measures.

For situational trust, a 3 (Capability of Automation) x 2 (Level of Understanding) x 4 (Event) mixed analysis mixed ANOVA will be performed to assess whether trust varied as a function of capability of automation across the three study drive events. Significant main effects will be examined via paired t-tests for separate dependent measures.

3. Understanding of Driving Automation Systems

- **How well do drivers understand driving automation systems after a short training, before they drive with them, and does this vary as a function of capability of automation and understanding manipulation?**
- **Does driving with a driving automation system increase understanding and does this vary as a function of capability of automation and understanding manipulation?**

Mean understanding scores will be calculated for the pre-drive questionnaire and post-drive questionnaire. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

A 3 (Capability of Automation) x 2 (Level of Understanding) x 2 (Time (pre vs. post)) mixed ANOVA will be performed to assess whether understanding varied as a function of capability of automation and understanding manipulation and whether there was a difference before and after the drive. Significant main effects will be examined via paired t-tests for separate dependent measures.

4. Risk Taking Propensity

- **Does risk taking propensity, measured by both self-ratings and objective assessment, relate to trust in driving automation systems?**
- **Is there a relationship between risk taking propensity, measured by both self-ratings and objective assessment, and takeover time, response time, and response severity?**
- **Is there a relationship between self-rated risk taking and objective risk-taking propensity?**

Mean self-rated risk-taking propensity scores will be calculated from the post-drive questionnaire. Mean number of pumps will be calculated across trials on the post-drive BART. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

Correlations will be calculated between risk-taking self-ratings and mean pre- and post-drive trust and situational trust to determine the relationship between self-rated risk-taking propensity and trust in driving automation systems.

Correlations will be calculated between mean pumps on the BART and mean pre- and post-drive trust and situational trust to determine the relationship between objective risk-taking propensity and trust in driving automation systems.

Correlations will be calculated between risk-taking self-ratings and mean pumps on the BART to determine the relationship between self-rated risk-taking propensity and objective risk taking.

Study 3

In Study 3, participants will drive the NADS-1 driving simulator with a driving automation system. Participants will be randomly assigned to one of six groups (Table 3), with different combinations of driving automation decision making (aggressive, moderate, conservative) and understanding (strong, weak). Automation decision making refers to how the automation behaves in the four different events in the study drive. Level of understanding refers to how well participants understand the functionality and limitations of the driving automation system and is manipulated via pre-drive training presentations.

Table 3. Study 3 Design

	Strong Understanding	**Weak Understanding**
Aggressive Automation	Group 1 (N=8)	Group 4 (N=8)
Moderate Automation	Group 2 (N=8)	Group 5 (N=8)
Conservative Automation	Group 3 (N=8)	Group 6 (N=8)

Study 3 will address the following research questions:

1. **Takeover Time, Response Time, and Severity**

- **Does takeover time differ based on the decision making of driving automation systems and level of understanding?**
- **Does response time differ based on the decision making of driving automation systems and level of understanding?**

Takeover time and brake and steering response time will be calculated for each study event. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

Data will be displayed descriptively, i.e., frequencies (%), mean, SD, and ranges. The predictor variables are those from sample characteristics, i.e., age, sex, gender, race, driving experience, trust in automation, risk-taking propensity, and mental model scores. The dependent variables relate to summary scores of the driving performance such as takeover time, response time, and severity. Variables will be assessed for normality to determine use of parametric versus non-parametric analyses via visual examination (e.g., probability plots, histograms, stem and leaf plots) and statistical tests (e.g., Fisher's skewness and kurtosis, Kolmogrov-Smirnov, or Shapiro-Wilks tests). Data will be disaggregated by sex and age. Inferential statistics capture differences between subjects understanding (strong vs. weak), level of automation (aggressive, moderate, and conservative) and within subject (*pre-drive vs post-drive*). Comparisons will be conducted via ANOVA. We will first use ANOVA, with a two-tailed test of significance ($p < .05$), to detect group differences and the Benjamini Hochberg (1995) procedure to identify the differences between levels of a factor, while controlling for multiple comparisons. Next, we will use linear models to regress the independent variables on the dependent variable(s) and to evaluate the effects of one independent variable on the dependent variable while controlling for other (confounding) variables. We will use log-linear models to regress the significant factors based on trust in automation and mental model scores, while controlling for confounders. Trend lines will be plotted to observe change scores for between subjects and within subject comparisons. We will use R (latest version) for data collation, mutation, and analysis.

2. Trust in Driving Automation Systems

- **How much do drivers trust driving automation systems before they drive with them and does this vary as a function of Automation Decision Making and level of understanding?**
- **How much trust do drivers have in driving automation systems across events as a function of Automation Decision Making and level of understanding?**
- **Does driving with a driving automation system increase or decrease trust and does this vary as a function of Automation Decision Making and level of understanding?**

Mean trust ratings will be calculated for the pre-drive questionnaire and post-drive questionnaire. Situational trust ratings will be captured for each study drive event. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

For pre- and post-drive trust, a 3 (Automation Decision Making) x 2 (Level of Understanding) x 2 (Time (pre vs. post)) mixed ANOVA will be performed to assess whether trust varied as a function of automation decision making level before and after the drive. Significant main effects will be examined via paired t-tests for separate dependent measures.

For situational trust, a 3 (Automation Decision Making) x 2 (Level of Understanding) x 4 (Event) mixed analysis mixed ANOVA will be performed to assess whether trust varied as a function of automation decision making level across the three study drive events. Significant main effects will be examined via paired t-tests for separate dependent measures.

3. Understanding of Driving Automation Systems

- **How well do drivers understand driving automation systems after a short training, before they drive with them, and does this vary as a function of capability of automation and understanding manipulation?**
- **Does driving with a driving automation system increase understanding and does this vary as a function of capability of automation and understanding manipulation?**

Mean understanding scores will be calculated for the pre-drive questionnaire and post-drive questionnaire. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

A 3 (Automation Decision Making) x 2 (Level of Understanding) x 2 (Time (pre vs. post)) mixed ANOVA will be performed to assess whether understanding varied as a function of automation decision making level and understanding manipulation and whether there was a difference before and after the drive. Significant main effects will be examined via paired t-tests for separate dependent measures.

4. **Risk Taking Propensity**

- **Does risk taking propensity, measured by both self-ratings and objective assessment, relate to trust in driving automation systems?**
- **Is there a relationship between risk taking propensity, measured by both self-ratings and objective assessment, and takeover time, response time, and response severity?**
- **Is there a relationship between self-rated risk taking and objective risk-taking propensity?**

Mean self-rated risk-taking propensity scores will be calculated from the post-drive questionnaire. Mean number of pumps will be calculated across trials on the post-drive BART. Data will be aggregated for all participants and the mean, median, and standard deviation for each dependent variable will be calculated.

Correlations will be calculated between risk-taking self-ratings and mean pre- and post-drive trust and situational trust to determine the relationship between self-rated risk-taking propensity and trust in driving automation systems.

Correlations will be calculated between mean pumps on the BART and mean pre- and post-drive trust and situational trust to determine the relationship between objective risk-taking propensity and trust in driving automation systems.

Correlations will be calculated between risk-taking self-ratings and mean pumps on the BART to determine the relationship between self-rated risk-taking propensity and objective risk taking.

5. <u>Provide the name and telephone number of individuals who were consulted on statistical aspects of the IC and who will actually collect and/or analyze the information</u>.

John Gaspar, PhD, University of Iowa, 319-335-4776

Justin Mason, PhD, University of Iowa, 319-467-1614

Chris Schwarz, PhD, University of Iowa, 319-335-4642

Eric Traube, NHTSA, 202-366-5673