Evaluation of Robot-to-Human Communication Designs Sampling Plan for Human Subject Recruitment

The Protective Technology Branch (PTB) of the Division of Safety Research (DSR) of NIOSH is conducting a project titled "Evaluation of Robot-to-Human Communication Designs." This project aims to investigate the effectiveness of robot-to-human communication designs. Data will be collected in NIOSH's Robotics Lab located in Morgantown, WV. The purpose of this Sampling Plan is to ensure, to the extent possible, that the participants will reflect the demographic composition of the Manufacturing workforce across the United States.

1. Respondent Universe and Sampling Methods

Potential Respondent Universe

This project includes data collection with manufacturing workers. The project aims to recruit a study population that is representative to the current composition of the manufacturing workforce in the United States. This composition, according to data from the Bureau of Labor Statistics (2023) [1], is 30% women, 78% White, 11% Black or African American, 8.0% Asian, and 18% Hispanic or Latino. Table 1 shows the potential respondent universe in the recruiting area.

Table 1. Potential Respondent Universe in Recruiting Area

Number of Manufacturing Companies in Recruiting Area [2, 3]	**Number of Manufacturing Employees in Recruiting Area [4, 5]**
There are approximately 1,906 manufacturing companies in the Pittsburgh Metropolitan Area.	86,300 individuals employed in manufacturing occupations.
There are approximately 151 manufacturing companies in the Morgantown Metropolitan Area	1,800 individuals employed in manufacturing occupations
Potential establishment respondent: 2,057	Potential workforce respondent population: 88,100

2. Sampling Methods

This study utilizes both a convenience and purposive sampling strategy. A convenience method is used based on organizations or individuals who respond to recruitment materials (Attachment A). From experience gained while conducting similar projects in the Virtual Reality Laboratory of NIOSH in Morgantown, West Virginia, one of the main challenges that researchers expect to encounter during the recruitment process is low response numbers. The reasons being that 1) the study

population has specific age (18 years of older) and experience (in manufacturing) requirements, i.e., Purposive Sampling, and 2) the relatively small population of the Morgantown metropolitan area. To circumvent these challenges, the recruitment area was extended to include the Pittsburgh metropolitan area.

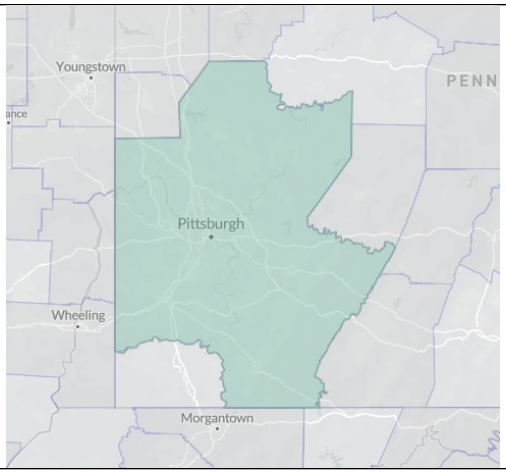
Initially, researchers will include all interested human subjects that meet the inclusion criteria, reflecting a random selection process of participants from a convenience sample. Then, as additional screening and eligibility processes occur (e.g., do not suffer from motion sickness), researchers will analyze the demographic composition of the human subjects that are volunteering to participate and will compare this composition against the targeted demographic composition aimed by the study, until the quota for every demographic composition is filled as much as reasonably practical, composing our purposive sample.

3. Respondent Selection Methods and Anticipated Sample

According to the Bureau of Labor Statistics, as shown in Figure 1, the demographic composition of manufacturing workings in the United States in 2023 was 30% women, 78% White, 11% Black or African American, 8.0% Asian, and 18% Hispanic or Latino. According to the CensusReporter website [6, 7], the metropolitan area of Morgantown, WV is shown in Figure 1a, and has a total population of 141,041 inhabitants, of which 47.9% are Women, 86.2% are White, 3.7% are Black or African American, 2.6% are Asian, and 2.8% are Hispanic or Latino. To increase the pool of potential participants, project researchers will expand the recruitment area to the Pittsburgh metropolitan area, which includes Pittsburgh, New Castle, Weirton PA-OH-WV combined statistical areas. The Pittsburgh metropolitan area, as shown in Figure 1, and Figure 1b, has a demographic composition of 50.8% Women, 82.7% White, 7.3% Black or African American, 2.7% Asian, and 2.1% Hispanic or Latino.

	Total Population	Percent Women (%)	Demographic Composition (%)			
			White	Black or African American	Asian	Hispanic or Latino
US Manufacturing Workforce	15,570,000	30	78	11	8	18
Morgantown, WV	141,041	47.9	86.2	3.7	2.6	2.8
Pittsburgh, PA	2,349,172	50.8	82.7	7.3	2.7	2.1

(a) Morgantown metropolitan area.



(b) Pittsburgh metropolitan area.

Figure 1. *Potential metropolitan areas for recruitment of human subjects.*

The target number of participants is 53. Table 2 represents the quota we will attempt to recruit which includes 16 females, 36 participants who are White, 6 that are Black or African American, 4 that are Asian, and 7 that are Hispanic or Latino. The race/ethnicity list is incomplete and people whose ethnicity is identified as Hispanic or Latino may be of any race, therefore the percentages don't add up to 100%.

Table 2. Quota for participants

Race/Ethnicity	**Manufacturing Percentage**	**Participant Sample**
Females	30%	16
White	78%	36
Black or African American	11%	6
Asian	8%	4
Hispanic or Latino	18%	7

4. Sampling and Recruitment Procedures

The instrument that will be used to recruit potential human subjects is a Recruitment Flyer, shown in Attachment A. Researchers will contact manufacturing companies and ask them to post the flyers in common areas of their workplace. Interested parties who call the telephone number listed on the recruitment flyer will be provided information about the project, its purpose, duration, and compensation; they will be asked two questions to determine whether they are susceptible to motion sickness, and if they meet the inclusion criteria, they will be e-mailed their scheduled date and time, along with a copy of the Informed Consent Form (Attachment B). The recruitment flyer will also be shared with the student and employee population at West Virginia University and posted in social networks like Facebook, and LinkedIn, to reach as many potential participants as possible.

5. Statistical Method for Sample Selection

The estimate for this sample (n=53) was calculated using GPower's repeated measures ANOVA sample size estimator. A small effect size of 0.10 power of 0.80, and a significance level of 0.05 were utilized to obtain this estimate. A mean dropout rate of 15.6% was observed in a meta-analysis of 46 studies that utilized head-mounted displays in virtual reality settings (HMD-VR) (Saredakis et al., 2020). We anticipate a similar dropout rate of around 15% for the current study. Therefore, up to 61 participants may be recruited.

6. Data Collection and Organization

This research study will take place at the NIOSH Research Robotics' Lab in Morgantown, WV. The total duration of the experiment session is approximately 4 hours. This study will use a Head-Mounted Display (HMD) virtual reality (VR) headset (HTC Vive Pro) to provide participants with a simulated sensory experience of the work environment. Used in conjunction with the HTC Vive Pro HMD headset will be the Vive VR controllers to simulate the human-robot collaboration task. The HMD system will be used to record the participant's reaction time and accuracy in responding to the 12 robot errors, and total time to task completion for each of the study conditions.

Following the HMD VR device setup, participants will take part in a 5-minute practice session to: 1) explore the virtual environment; 2) practice picking up and releasing objects as the robot passes them to the participant; and 3) practice responding to robot error conditions by pressing one of the three buttons on the virtual response panel.

Participants will perform the task while standing stationary in the corner of the NIOSH Robotics Lab. The task requires only upper body movement. The area is free of obstructions and distractions. Two NIOSH researchers will be present to observe and instruct the participant while they participate. Researchers will also observe and monitor the participant to watch for risks such as potential loss of balance and intervene with the participant if necessary.

Participants will progress through and complete 20 simulated conditions (5 modalities x 2 message types x 2 safety-relevancies). In the Robotics Lab, which has a simulated VR environment, participants will be tasked with receiving objects from the virtual robot as the robot arm moves toward them. There will be three types of errors encountered during the simulations: gripper, arm, and body errors. Each error type will occur 4 times in each condition.

The errors will occur randomly throughout each condition, and the order of different types of errors will be randomized. During the error trial, the robot will exhibit a sign

and move at a slowed speed. When an error occurs, the robot will communicate the error via the specific interface modality and message type of the current condition. The participant then has a limited amount of time to respond to the error on the response panel. Within each condition, data will be collected for each participant's time to complete the condition and response accuracy and reaction time. After each condition, participants will respond to questions regarding their perception of safety and trust in the robot, difficulty of the interface modality, mental and team workload, and their evaluation of the robot. Participants will be given a short break after every three conditions or per request.

After completion of the study trials, researchers will remove the HMD-VR headset from the participant and commence the debriefing. Researchers will check participants' level of nausea using the Virtual Reality Sickness Questionnaire and pay them for their time via a debit card ($30/per hour). If the participant is experiencing any level of nausea based on their responses to the questionnaire, they will not be released until symptoms subside.

Description of Human-Robot Interaction Conditions

1. **Interface modality:**
    a. *Visual-lights.* This modality is most often seen in collaborative robots and considered as a control condition.
        i. Normal Operation: The robot will display a solid green light (this is a normal LED light encountered in everyday settings – like a green light on a working smoke alarm).
    b. Visual-graphical
        i. Normal Operation: The robot will display a green checkmark on the screen (e.g., see Figure 2).
    c. Visual-verbal
        i. Normal Operation: The robot screen will display text that says "Normal Operation."
    d. Auditory-tone
        i. Normal Operation: There will be a beeping sound at 500 Hz and 200 ms duration. (500 Hz is the lowest frequency to which the human auditory system is most sensitive [between 500 and 4000 Hz]).
    e. Auditory-speech
        i. Normal Operation: Same as the auditory-tone interface modality. There will be a beeping sound at 500 Hz and 200 ms duration.

2. **Message type:**
    a. *Status-based:* conveys the current status of the robot's operation
    b. *Intention-based:* conveys information indicating [future] intention of the robot

3. **Safety-relevancies**:
    a. *Low:* simulated non-hazardous object is being handled (e.g., smooth object)
    b. *High*: simulated hazardous object is being handled (e.g., sharp object)

Description of Dependent Variables

1) Task Performance
    a. Accuracy of a user's understanding of robot messages (measured by selection of error location)
    b. Attention toward the communication to a given interface (measured by the reaction time)
    c. Time to task completion

2) **Self-reported measures**:
    a. *Perceived Difficulty of Interface Modality*: 1 statement indicating the degree of difficulty the participant experienced in understanding the messages from the robot on a 7-point scale from "very difficult" to "very easy."
    b. *Perceived Safety*: A statement indicating the participants' feelings of safety while working with the robot on a 7-point scale from "completely disagree" to completely agree."
    c. *Perceived Trust*: A statement indicating the participants' feelings of trust while working with the robot on a 7-point scale from "completely disagree" to "completely agree."
    d. *Team Workload*: 9 questions assessing individual and team workload that have been modified to fit the collaborative task; rated on a scale from 0 (minimal) to 10 (maximum).
    e. *Robot Evaluation*: 14 questions rated from 0 to 100% that evaluate different aspects of robot behavior and performance.

Data analysis

A series of repeated measures ANOVA will be used to examine the effects of different interface modalities, message types, and safety-relevancies on the accuracy, reaction time, time to complete the task, and self-reported measures. Two-tailed, 0.05 p-value, and 95% confidence intervals will be used for statistical significance. If the sample size permits, correlational and regression analysis may be used to further explore the relationships among variables. Tables and figures will be generated to illustrate the differences between the conditions and any interaction effects.

7. Methods to Maximize Response Rates and Deal with Nonresponse

To maximize participation, we will offer the following incentives: 1) Compensation for the participants time at a rate of $30/hour, with partial payments at a rate of $10 per 20 minutes, with periods less than 20 minutes rounded up; 2) flexible hours that include early hours in the mornings and/or evening hours. In addition, to boost the recruitment rate, we will provide all human subjects who participate information about the study to share with their coworkers, if they desire.

To deal with nonresponse or participant drop out rates, the team intends to overrecruit. As described earlier, a mean dropout rate of 15.6% was observed in a meta-analysis of 46 studies that utilized head-mounted displays in virtual reality settings (HMD-VR) (Saredakis et al., 2020). We anticipate a similar dropout rate of around 15% for the current study. Therefore, up to 61 participants may be recruited.

References

[1] Bureau of Labor Statistics (BLS). 2023. Labor Force Statistics from the Current Population Survey. Employed persons by detailed industry, sex, race, and Hispanic or Latino ethnicity : U.S. Bureau of Labor Statistics

[2] Dun and Bradstreet 2024 – Find Manufacturing Companies in Morgantown - Dun & Bradstreet

[3] Dun and Bradstreet 2024 - Find Manufacturing Companies in Pittsburgh - Dun & Bradstreet

[4] Bureau of Labor Statistics (BLS). 2024. Economy at a Glance (Morgantown, WV). Morgantown, WV Economy at a Glance

[5] Bureau of Labor Statistics (BLS). 2024. Economy at a Glance (Pittsburgh, PA). Pittsburgh, PA Economy at a Glance

[6] Pittsburgh-New Castle-Weirton, PA-OH-WV CSA - Profile data - Census Reporter

[7] Morgantown-Fairmont, WV CSA - Profile data - Census Reporter

[8] Saredakis, D., Szpak, A., Birckhead, B., Keage, H. A., Rizzo, A., & Loetscher, T. (2020). Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis. Frontiers in Human Neuroscience, 14.