User Study Design Example

Purpose of Study
- What do you expect to get out of the study?
  - Do people react differently to a robot that:
    - Gives spoken responses as it follows instructions (the speaking case).
    - Follows instructions silently (the silent case).
  - Whether the preferences and reactions differ when the robot makes mistakes.
  - Whether these conditions change how long the participant is willing to engage with the robot.
- What tasks will the participants perform?
  - They will give the robot a sequence of predefined instructions, such as “pick up the blue block.”
  - They will be offered a chance to stop after five instructions, and the task will end after ten instructions.

Hypotheses
- What are your hypotheses?
  - Hypothesis 1: Participants will be more comfortable with a robot that gives a sign that it has heard the instruction, e.g., “Okay, picking up the blue block.”
  - Hypothesis 2: Participants will find the “silent” case more acceptable when the robot misinterprets a command, because not being able to correct the robot will be more annoying.
  - Hypothesis 3: Participants will spend more time with the speaking robot.

Participants and Setup
- How many people do you need for the study?
  - We need fifteen participants total; 20 would be better.
- How will you get them?
  - Ask classmates; ask friends; post flyers.
  - We will start posting flyers and scheduling participants on April 10th, and will send our flyer to the professor by April 8th.
- How will you instruct the participants?
  - We will read them a short description of the task (drawn from this overview), then show a sample video of ourselves reading off two instructions, with the robot’s responses shown.
  - No familiarization tasks.

Metrics
- What are your dependent variables?
  - The participants’ level of comfort, level of frustration, impatience, and level of willingness to use the robot.
- What will you measure?
  - Participants’ self-reported levels of comfort and frustration.
  - Whether the user asks to stop after 5 trials.
- How will you measure that?
  - Questionnaires for emotional state reactions.
  - Experimenter tracking of when the user chooses to stop.

Instruments
- What will you ask in your questionnaire?
  1. 5-point Likert scale agree/disagree questions for each of roughly 8 emotions, for example, E.g.: I found the robot frustrating.
  2. 5-point Likert scale agree/disagree questions comparing the two cases. E.g.: I preferred it when the robot told me what it was going to do.
  3. Short (two-line) questions about the subjective experience. E.g.: “What did you find most appealing?” E.g.: “If you chose to stop after five interactions, why?”
  4. We will also track how often the robot crashes for each user.
- Will you have a pre, post, or per-task questionnaire?
  - (1) will be pre-task; (2) and (3) will be post-task.
Materials

- Which robot platform and peripherals are you using?
  - Both Jaco arms; a (new) Kinect; a webcam for the “wizard” to use while observing from another room; computer speakers.
- What capabilities will be used for your study?
  - Text-to-speech, moving to an arbitrary point in space, manipulating (picking up) objects.
- What props (tables, objects, obstacles, etc.) do you need?
  - The arm, the table it’s attached to, brightly colored blocks.
- Will you use actors? What exactly will the actors do?
  - I will not use actors.

Autonomy

- What will be autonomous? Wizard-of-Oz!
  - The experiment will be part WoZ, part autonomous action.
  - The wizard (me) will generate speech by pressing keys for predefined phrases the robot will say and what block to pick up.
  - The robot will autonomously find colored blocks on the table, using the Kinect, and pick them up using the arm.
  - My partner will remain in the room, but not engage with the experiment except to restart the robot in case of crashes.

Study Conditions

- What kind of study is it?
  - This will be a comparative, within-participant study.
  - Each participant will be exposed to the silent and the speaking cases, in random order.
  - Randomizing the order will average out task fatigue effects.
  - They will be asked to comment on each case, then afterwards, to compare the two.
  - This will control for the timing frustration involved in any interaction with a robot.
- Where will you conduct the study?
  - The study will be carried out in ITE 343.

Example Interaction

1. Intro: experimenter reads the instructions, gathers the signed consent form, shows the video of the sample interactions, and answers questions.
2. Participant (reading from a whiteboard behind the robot): “Please pick up the red block.”
3. Wizard hits the key for response 1:
   - Jaco arm: “Picking up the red block.”
   - <moves to pick up the first block, drops red block in a bin, returns to resting state>
4. Participant (reading): “Put away the blue block.”
5. Wizard hits the key for response 2:
   - Jaco arm: “Picking up the green block.”
   - <moves to pick up the wrong block, drops green block in a bin, returns to resting state>
   - [participant goes through three more cases]
6. Experimenter: “That’s five instructions. Would you be willing to do five more? Any answer is fine.”
7. Participant: “I think I’ll stop now.”