Carnegie Mellon University

Exempt Review IRB Application/Protocol

Please complete this application as thoroughly as possible. Give as much detail as possible. If a question is not applicable to your study, please indicate this by responding "N/A".

Additional information and templates are available at Office of Research Integrity and Compliance
*Note that incomplete applications will result in delays.

Protocol Number: (must match the study number provided by SPARCS)

STUDY2023_00000036

Study Title: (must match title in SPARCS form)

Large-Language-Model-Supported Debugging and Testing Practices for CS Learners (Pilot)

PI Name:

Tongshuang Wu

1. Exempt Category(ies)

Select the Exempt Category(ies) that describe your research.

☐ **Category 1 - Protocols that are conducted in established educational settings and involve normal educational practices that are not likely to adversely impact students' opportunity to learn required educational content or the assessment of educators who provide instruction.**

☐ **Category 2 - Protocols that only involve surveys or interviews of adults as well as educational tests or observation of public behavior**

One of the following criteria must be met (select at least one):

☐ **The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects;**

☐ **Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation;**

☐ **The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review.**

☒ **Category 3 - Research involving benign behavioral interventions in conjunction with the collection of information from an adult subject through verbal or written responses (including data entry) or audiovisual recording if the subject prospectively agrees to the intervention and information collection. Benign behavioral interventions are brief in duration, harmless, painless, not physically invasive, not likely to have a significant adverse lasting impact on the subjects, and the investigator has no reason to think the subjects will find the interventions offensive or embarrassing.**

One of the following criteria must be met (select at least one):

☐ **The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects;**

☒ **Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation;**

☒ **The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review.**

☐ **Category 4 - Protocols that only involve Secondary Use without Consent meaning the collection/study of existing (in existence at the time of submission) data, documents, records, or specimens if all are publicly available or recorded such that individuals cannot be identified**

☐ **Category 5 - Research and demonstration projects that are conducted or supported by a Federal department or agency, or otherwise subject to the approval of department or agency heads (or the approval of the heads of bureaus or other subordinate agencies that have been delegated authority to conduct the research and demonstration projects), and that are designed to study, evaluate, improve, or otherwise examine public benefit or service programs. Please contact the IRB at irb-review@andrew.cmu.edu prior to IRB submission if you think your study qualifies for this exemption.**

☐ **Category 6 - This category is limited to taste and food quality evaluation and consumer acceptance studies (i) If wholesome foods without additives are consumed, or (ii) If a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or**

environmental contaminant at or below the level found to be safe, by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the U.S. Department of Agriculture. Please contact the IRB at irb-review@andrew.cmu.edu prior to IRB submission if you think your study qualifies for this exemption.

If the research involves other activities, it is not eligible for Exemption. Do not proceed. Return to SPARCS to the Review Type Requested page to change to a Non-Exempt application and use the Non-Exempt Protocol Word document.

2. Study Scope

a. What is the purpose of the study (what is your research question) and how will the data collected be used?

This study will examine how to use Large Language Model-based AI code-generation tool to teach hypothesis-testing skills to programmers (undergraduate students or teaching assistants in a programming course at CMU, e.g., 15110 Principles of Computing and 15112 Fundamentals of Programming).
The data would be collected through think-aloud and surveys, during which the participants will be using our tool. The clickstream data will also be collected while the participant is interacting with the tool and used by the researcher to analyze the participants' debugging and testing process and progress.
This IRB protocol is a pilot study of the full study, as we aim to collect some think-aloud data to finalize the design of our tool and survey questions.

b. For each activity/participant population, describe the research procedures. In the SPARCS system, provide any questionnaires, surveys, tools, device manuals, Manual of Procedures, protocols, etc. that will be used to collect data or to direct the conduct of the study.

After acquiring participants' consent before the study via online consent form, in the study session, participants will first fill out a Google form (based on the CMU Google suite) with demographic information and prior experience in programming. TA and students will be recruited but they will complete the activity individually and will not be matched.

Then, participants (students or TAs) will be provided with a tutorial (sample interface demonstration slides attached) on how to interact with the tool they'll use for the practice and a demonstration of the think-aloud technique (sample script attached).

During the practice with the tool, participants (students or TAs) will see the homework problem they have done and develop a test suite with the goal of detecting bugs in code (a sample question is used in the interface demo slides, the actual problems being used will be the actual homework problem of this class, which is still under development). Think-aloud notes will be taken.

Finally, participants will fill out another google form survey asking about their interaction experiences and ask the students to self-report their grade and score and submit their solutions to the homework, the TAs will be asked to provide the feedback they have given students in the homework. the researcher may ask the participant to explain their answers and take notes.

c. For each activity/participant population, indicate the location(s). Specify whether the participant will be engaged in person, remotely via the internet, etc.:

Participants will complete the think-aloud session in-person with the researcher in booked conference rooms in Newell-Simon Hall or the Gates and Hillman Centers.

d. For each activity/participant population, describe the time required of the participant (time for each study visit AND overall time commitment for the study):

One think-aloud session will be approximately 30-40 minutes for a participant. Each participant will complete at least 2 problems in the tool, which take experts approximately 5 minutes each, and may take a participant about 10 minutes. We consider extra time for surveys and tool familiarization.

e. Will questionnaires or surveys be used?

If yes, please upload to the Study-Related Documents section of the SPARCS application.

3. Participation Information

a. What is the age range of participants in the proposed study?

≥ 18

b. How many participants/records/specimens are needed for the study? Cannot be "unlimited"

Up to 10 participants.

c. How was that number determined? Provide power analysis or other justification for the number requested.

This think-aloud is only a pilot study of the full study. As we aim to collect some think-aloud data to finalize the design of our tool and survey questions, we estimate to recruit 3~5 participants each from programming courses like 15110 and 15112 to participate in the think-aloud activity.

d. Please list <u>all</u> inclusion and exclusion criteria for your selection of subjects:

Participation in this study is limited to CMU undergraduates aged 18 and older currently a student or a teaching assistant (TA) in a programming class such as 15110 or 15112.

e. What do you estimate the ratio of males to females to be?

Male : Female ≈ 1:1 given the enrollment & TA statistics of a programming course like 15110 & 15112

f. Will this be reflective of the local population?

If not, please explain:

g. Will vulnerable subjects (Pregnant Women, Neonates, Prisoners, Children, and Cognitively Impaired Adults) be involved in the proposed study?

- **Pregnant women, human fetuses?**
- **Neonates?**
- **Prisoners?**
- **Children?**
- **Cognitively Impaired Adults?**

h. Will you target a certain population?

Please explain:

We target participants who're currently taking or TAing a programming class like 15110 or 15112 this semester.

i. Do you anticipate that your participants will represent a cross-section of the population in the region where the study is being conducted?

If no, please describe your study population and address why minority representation is not considered:

We work with students enrolled at CMU and currently taking or TAing the introductory programming course. This study population can be expected to roughly represent the proportion of minorities at CMU, but only roughly because not all students of CMU need to take/TA a programming class like 15110 or 15112.

j. Will subjects from non-US locations be enrolled?

If yes, please describe how you will ensure that the research will not violate the cultural norms and values of the subjects or put the subjects at risk not normally encountered by US subjects: Please note that in some cases, a cultural evaluation may need to be completed by an impartial third party.

If yes, please provide information below:

Name of Non-US Country(ies) from which subjects will be enrolled	Number of subjects to be enrolled from this Non-US Country
Click or tap here to enter text.	Click or tap here to enter text.
Click or tap here to enter text.	Click or tap here to enter text.
Click or tap here to enter text.	Click or tap here to enter text.
Click or tap here to enter text.	Click or tap here to enter text.
Click or tap here to enter text.	Click or tap here to enter text.
Click or tap here to enter text.	Click or tap here to enter text.

4. Recruitment

a. Describe how participant recruitment will be performed:

Participants will be recruited via an announcement made to the class through piazza (the course Q&A platform used in a programming course like 15110 & 15112) with the help of instructors.

b. Indicate how and by whom potential participants are introduced to the study:

Participants will be introduced to the study by the investigator of this project as well as the instructor of this class through course announcements.

c. Check all boxes below that apply and attach documentation:

☐ **Flyers? Where will they be posted?**

☐ **Radio, TV?**

☒ **E-Mail? Indicate how the email addresses are obtained:**

With the help of instructors, and via the piazza messaging platform, which automatically sends an email to participants if it's an instructor's announcement.

☐ **Web-based? – NOTE: If recruiting on mTurk or another similar online system, the title of the HIT/advertisement must include the SPARCS study number.**

☐ **Participant Pool? Which one?**

☐ **Other? Describe:**

d. Will participants undergo screening prior to their participation?

If yes, please describe:

Please upload all recruiting and screening materials in the Study-Related Documents Section of the SPARCS application.

5. Consent

Describe the process for providing subjects with information about participation and confirming their willingness to participate:

Consent will be gathered before the participant schedule a data collection session via voluntary submission of the online consent Google form.

Please provide any Informed Consent Forms, introductory scripts, or other documents that you will use to inform subjects (or parents, if subjects are minors) about the research. (Examples include a script to be used to explain the study verbally, an information sheet to be given to subjects, an online consent page, etc.) These documents should be uploaded in SPARCS.

6. Risk and Benefits

a. **Will participants receive a direct benefit from the study?**

If yes, describe the expected direct benefits to participants:

b. Indicate the expected indirect benefits to participants/scientific knowledge:

We expect participants to get some educational values of debugging & testing practices from interacting with our tool.

c. Indicate all potential risks to participants:

The risks and discomfort associated with participation in this study are no greater than those ordinarily encountered in daily life. Participants struggling to complete the debugging & testing practice may feel frustrated or unsatisfied with their performance.

d. Indicate how all potential risks will be managed and/or minimized:

Participants will be reminded that it is okay to not finish all exercises (as it's meant to be practice that have no influence on their grades) and that in fact they are designed to challenge them so that they should just give it their best try.

7. Deception (Exempt Category 3 Only)

Deception is only possible in minimal risk studies. Investigators need to explain why the deception is necessary to achieve the study goals and how the degree of deception is kept to a minimum. The degree of deception means, for example, withholding part of the study's purpose as opposed to stating a false study purpose. Subjects should be debriefed as early as is feasible. If the research involves deceiving the subjects regarding the nature or purposes of the research, this exempt category is not applicable unless the subject authorizes the deception through a prospective agreement to participate in research in circumstances in which the subject is informed that he or she will be unaware of or misled regarding the nature or purposes of the research.

a. Will deception be used?

If no, proceed to the next section.

If yes, please explain:

If yes, please include a justification as to why deception is necessary:

b. Indicate how you will ensure that the subjects authorize the deception through a prospective agreement to participate in research in circumstances in which the subject is informed that he or she will be unaware of or misled regarding the nature or purposes of the research:

c. **Describe how participants will be debriefed:**

Please upload the de-briefing material and/or script in Study-Related Documents section in the SPARCS application.

8. Compensation

a. **Are participants to be compensated for their study participation in any way?**

If no, proceed to the next section.
If yes, what is the amount of compensation?

If yes, what is the source of the compensation?

If yes, what is the type of compensation (eg, gift card, cash)?

b. **Will participants receive any non-monetary compensation (e.g., parking validation, snacks, lottery tickets)?**

If yes, please describe:

c. **Are there any costs to participants?**

If yes, please describe:

d. **Will participants who are students be offered class credit?**

9. Data Security and Confidentiality

a. Type(s) of data to be collected/accessed/obtained/used for this research (select all that apply).

☐ **De-identified (ALL identifiable information is completely and permanently removed with NO code or link to the original identifiers. The CMU researchers were never able to identify the subjects. Data can never be re-identified.)**

☐ **Coded data (all identifiable information is completely removed from the data set, but a code or link exists enabling the possibility of re-identification of subjects)**

- **Will the code or link be accessible/available to the research team? If yes, data is identifiable.**

☐ **Anonymous data (no identifiable information was/will be collected or obtained, no identifiable data ever existed)**

☒ **Identifiable data (research contains any elements that allow for the identification of a subject)**

If identifiable, provide a complete list of identifiers:

Email address (Andrew id)

b. Describe the information being collected and/or used for the research. Provide a list of all data elements to be included in the research. You may provide a list here OR upload a spreadsheet or list of all data points into the SPARCS application in the Study-Related Documents page.

- Demographics (Age, gender, major, year at CMU)
- Self-reported prior experience & comfort level with programming, debugging & testing
- Grade/performance in class & assignment (i.e., students' current grade in the course, the score they received for the specific homework problem they practiced in the tool)
- Assignment artifacts (i.e., students' homework submissions for a specific homework that include the practice problems)
- TA feedback (i.e., the feedback participants (students) received, or the feedback the participants (TA) gave for the practice problem)
- Tool interaction clickstream data (time-stamped log data)

c. Will the research use existing data sets/recordings/specimens?

If yes, describe: 1) where the data/recordings/specimens originated, 2) whether they were collected with consent from the subjects, and 3) if they were collected under a different IRB protocol. Provide any related IRB protocol numbers below and upload any related informed consents in the Study-Related Documents page in the SPARCS application.

d. Describe your procedure for coding your data (encoding):

Log data (time and clicks) of their interaction with the tool will be collected on a secure server, and no IP address or any identifiable information other than their Andrew id/email will be included. Students' grades information will be obtained from the students (self-report), and stored on password-protected, secure cloud services (e.g., CMU's G Suite or CMU's Box). All Andrew id & emails will be removed & replaced with an anonymous subject ID once all data collection is completed, so only the de-identified data will be used for the analysis, and the original data containing identifiable information will be securely deleted.

e. **Will audio recordings be made?**

If yes, please describe how this will be done and who will have access to the recordings:

f. **Will video recordings be made?**

If yes, please describe how this will be done and who will have access to the recordings:

g. **Do you intend to obtain a certificate of confidentiality from NIH?**

h. <u>In addition</u> to the individuals listed on the study personnel page, who will have access to research data (e.g. surveys, questionnaires, recordings, interview records, etc.)? Include a comprehensive list and indicate if information may be shared outside the research team and/or CMU (including collaborators, vendors, sponsors, etc.). Include what data each party will have access to and how the data will be transmitted/shared with them.

The study data will be analyzed by the research team only.
Only the research team member will have access to the raw & identifiable research data, and all identifiable information will be removed once data collection is completed, and then the research team members will only hold onto the de-identified data and perform data analysis. In the future we may share the de-identified, aggregated data & analysis with other researchers.

i. Describe how you will protect participant confidentiality and secure research records (e.g. password protected, encrypted, etc.). Include location of where the data will be stored. If the PI should leave the university indicate your plan for the storage of research information and who will be responsible for oversight.

To protect participants' identities during this study: (1) Each participant will be assigned a subject ID number, which will be used to anonymize the participants (so the emails and Andrew ids will be replaced with the de-identified subject ID); (2) Any original log data or assignment grades & artifacts containing any personally identifiable information will be stored in a secured location (secure server or password-protected laptop, with a weekly secure CMU cloud-services backup) accessed only by authorized researchers of this research team. Once the de-identified process is completed, the original data including identifiable information will be securely removed.
<u>If the PI leaves the university, Dr. Ken Koedinger (Faculty Advisor) would be responsible for the storage and oversight of the data at CMU.</u>

j. Describe your process for overseeing your study. Include a description regarding monitoring of data (to ensure that study goals are met and adherence to the IRB approved protocol is maintained). Examples: Review of lab notebooks, frequency of meetings to review data, who will be present at the meetings, how recruitment and retention will be monitored, etc.:

The PI will be regularly updated (usually weekly) by the members of the research team regarding data collection, data anonymization, and data analysis.

k. Describe your process for ensuring that adverse events, unanticipated problems, and subject complaints are reported to the IRB Office in a timely manner: Please note, all reportable events are required to be reported to the IRB via a Reportable New Information (RNI) form in SPARCS within 3 days.

The PI will react quickly when receiving questions or comments regarding the study, or when the regular updates from members of the research team mention signs of possible trouble. Adverse events and unanticipated problems will be reported promptly to the IRB Office.

l. Please describe the intended final state of all data collected/obtained/used for this study (e.g., will you initially be collecting identifiable information for this research, but during the course of the study you will de-identify it?). Final state of data will be:

☒ De-identified (ALL identifiable information is completely and permanently removed with NO code or link to the original identifiers. Data can never be re-identified.)

☐ Coded data (all identifiable information is completely removed from the data set, but a code or link exists enabling the possibility of re-identification of subjects)

☐ Anonymous data (no identifiable information was collected or obtained, no identifiable data ever existed)

☐ Identifiable data (data contains any elements that allow for the identification of a subject)

m. **Confirm that all research data will be retained at CMU for a minimum of three (3) years past study completion: ☒**

10. Collaborators

a. Is this research to be done in collaboration with any institutions, individuals, or organizations not affiliated with CMU?

If no, STOP here.

b. List the collaborators involved in the study:

Names of Collaborating Institutions, Individuals, or Organizations	Is CMU overseeing this Collaborator? (may require a signed Reliance Agreement)	Is this Collaborator overseeing CMU? (may require a signed Reliance Agreement)	Will Collaborator participate in interaction or intervention with subjects?	Will collaborator receive/have access to identifiable data?	Will Collaborator consent subjects?
Click or tap here to enter text.	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No
Click or tap here to enter text.	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No
Click or tap here to enter text.	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No
Click or tap here to enter text.	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No
Click or tap here to enter text.	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No
Click or tap here to enter text.	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No	Enter Yes or No

c. Does CMU have responsibility for oversight of the entire study?

If no, describe the monitoring/Quality Assurance process for the entity(ies) who will have responsibility for oversight of the study:

d. Is there IRB approval from another IRB for this study?

If yes, attach the IRB approval documents to Question 2 on the Local Site Documents page in the SPARCS application.