IRB Protocol Application

Primary Info

Protocol Number: [To be assigned]

IRB Protocol Title: UX Metrics Impact on Designers' Decision-Making

Lay Summary: This study investigates how different sources of user experience evaluation data (AI-generated vs human-collected) and different metric types (standard UX vs ethics-enhanced metrics) influence professional designers' willingness to implement potentially manipulative interface designs. Participants will view interface mockups alongside evaluation data from both AI and human sources, presented with either standard UX metrics or ethics-enhanced metrics, and indicate whether they would recommend implementing these designs. The goal is to understand how AI evaluation tools compare to human evaluation in design decision-making and whether enhanced ethical metrics influence decisions regardless of evaluation source.

Is this a student project?: Yes

Type of project: PhD student research

Please describe: PhD student research spun out from prior published work on dark patterns and UX metrics.

Research Team

Principal Investigator: Mr. Hauke Gregor Wilhelm Sandhaus (hgs52)
Department/Center: Information Science
Business Title: Graduate Research Assistant
Edit Permission: YES

Faculty Advisor: Prof. Helen Nissenbaum (hn288)
Department/Center: Information Science
Business Title: Professor
Edit Permission: YES

Review Type Determination

Do you need proof of preliminary IRB approval for a funding agency before your study materials are fully developed and ready for complete IRB review?: No

Will an external IRB act as the IRB of record for this study?: No

Are other institutions engaged?: No

Does your project involve any of the following:

- FDA regulated drugs and devices: No

- Prisoners or other institutionalized individuals: No
- Active collection of biospecimens: No
- Conducting of biomedical procedures: No

Which of the following types of activities will your research involve?:

- Category 2: Research involving educational tests, survey or interview procedures, or observation of public behavior

---

Exempt Determination

Category 2: Research involving educational tests, survey or interview procedures, or observation of public behavior: Yes

Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met: (i) The information obtained is recorded by the investigator in such a manner that the identity of the human participants cannot be readily ascertained, directly or indirectly through identifiers linked to the participants; OR (ii) Any disclosure of human participants' responses outside the research would not reasonably place the participants at risk of criminal or civil liability or be damaging to the participants' financial standing, employability, educational advancement, or reputation.

Does your study involve the collection of or use of human derived materials such as blood, bodily fluids, stem cells, tissue, or fetal tissue?: No

Special Populations: None of the above

Auto-determined Review Type: Exempt

Protocol Description

Purpose and goals of the research

The purpose of this research is to examine how evaluation source (AI vs human) and metric type (standard vs ethics-enhanced) influence professional designers' decision-making when evaluating interface designs. This study uses a 2x2 experimental design:

- **Between-subjects factor**: Metric type (Standard UX metrics vs Enhanced Ethics metrics)
- **Within-subjects factor**: Evaluation source (AI-generated vs Human-generated evaluation data)

The study aims to determine: (1) whether AI evaluation data influences design decisions differently than human evaluation data, (2) whether enhanced ethical metrics affect decisions regardless of evaluation source, and (3) how designers perceive the credibility and usefulness of AI vs human interface evaluation.

How will this study contribute to existing knowledge?

This study will contribute to existing knowledge by providing empirical evidence on: (1) the comparative influence of AI vs human evaluation data on design decisions, (2) whether enhanced ethical metrics serve

as effective interventions across different evaluation sources, and (3) professional designers' trust and reliance on AI evaluation tools. As AI tools become more prevalent in design workflows, understanding their impact on ethical decision-making is crucial for developing responsible design practices.

Will the research involve secondary use of data, documents, records or biospecimens collected from individuals?

Yes - Human evaluation data from a previous study will be used.

Description of secondary data use: The human interface evaluation dataset consists of aggregated UX scores (efficiency, satisfaction, ease of use, clarity, and ethical metrics) collected from human participants in a previous study. Only aggregate averages per interface will be presented to participants - no individual-level data from the previous study will be used or disclosed.

The AI evaluation dataset was generated using OpenAI and Claude vision language model APIs to analyze interface examples, without any human participant involvement. These AI-generated scores will be presented alongside the human evaluation averages.

Data source and permissions: [Add details about the previous study - IRB approval number if applicable, permission to use data, etc.]

Describe what participants will be asked to do and/or what information or specimens will be collected from them.

Participants in this study will be asked to evaluate interface designs using data from both AI and human sources. The study will be conducted online through a survey platform. Participants will:

1. Complete a brief demographic questionnaire about their design experience, role, and prior AI tool usage
2. View a series of interface mockups (approximately 10-12 interfaces)
3. For each interface, view evaluation data from TWO sources:
   - Human evaluation data (aggregated scores from a previous study with 120 human participants)
   - AI evaluation data (generated using OpenAI and Claude vision language models)
4. Evaluation data presented as either:
   - Condition A: Standard UX metrics (efficiency, satisfaction, usability)
   - Condition B: Enhanced metrics including ethical considerations (manipulation, deception, addiction potential)
5. Answer decision questions for both evaluation sources: "Would you release this interface design based on [human/AI] evaluation data?"
6. Provide brief explanations for decisions based on each evaluation source
7. Answer comparative questions about which evaluation source was more influential
8. Complete follow-up questions about AI vs human evaluation preferences and decision-making process

Provide a sequential list of all study components, including any follow-up sessions. List the estimated time needed for each component, as well as the total time commitment for participants.

1. **Informed consent and eligibility screening** (2 minutes)

2. **Demographic and professional background questionnaire** (3 minutes)
3. **Interface evaluation task** (25-30 minutes)
   - Review 10-12 interface designs with both AI and human evaluation data
   - Make implementation decisions for each evaluation source
   - Provide reasoning for decisions
   - Answer comparative questions
4. **Post-study questionnaire about AI vs human evaluation and decision-making process** (7 minutes)
5. **Debrief and study completion** (2 minutes)

Total estimated time: 35-40 minutes

Will the research team input participant data into an AI tool or rely on software that makes use of AI features?

Some of the evaluation data shown to participants will be AI-generated (clearly labeled as such), but no participant data will be input into AI tools during the study. The AI evaluation data will be pre-generated and static.

Will this study involve participants directly interacting with an AI tool or service?

No

Describe any other websites, software, or electronic resources that will be used for data collection.

The study will be conducted using Qualtrics survey platform (https://cornell.yul1.qualtrics.com). Interface mockups will be hosted on a secure platform (likely Framer.com or similar) and embedded in the survey. Participants will be recruited through professional networks and possibly Prolific for supplemental recruitment.

Recruitment

Specify where the research will be conducted.

The study will be conducted entirely online through Cornell Qualtrics platform.

Anticipated start date of the research**: [To be determined based on IRB approval]

Describe the sampling plan and the sample size.

We plan to recruit 120-160 professional designers and design decision-makers to achieve adequate power for the 2x2 design (60-80 participants per between-subjects condition). Participants will be stratified by role and experience level. We will use convenience sampling through professional networks, design communities, and potentially Prolific for additional recruitment.

Which of the statements describes the recruitment strategy?

Potential participants will first learn of this research opportunity from a flyer, advertisement, email/phone call, presentation, or similar prospective recruitment strategy.

For prospective recruitment strategies, describe the recruitment procedures.

Recruitment will occur through:

1. Professional design communities and forums
2. Social media posts in design-related groups
3. Direct outreach to design professionals through LinkedIn and professional networks
4. Potentially Prolific platform for supplemental recruitment
5. Snowball sampling through initial participants

Indicate the total number of participants to be recruited.

160

List the inclusion and exclusion criteria.

Inclusion criteria:

- Professional experience in UI/UX design, product design, or design decision-making roles
- Minimum 1 year of professional design experience
- Age 18 or older
- English proficiency sufficient to complete survey

Exclusion criteria:

- No professional design experience
- Students without professional experience
- Under 18 years of age

Does your research involve recruitment of participants in any of the following countries/regions: Cuba, Iran, North Korea, Sudan, Syria, Venezuela, Ukraine, Russia or Belarus?

No

Describe the eligibility screening process.

Participants will be screened through initial survey questions:

1. "Do you have professional experience in UI/UX design, product design, or design decision-making?"
2. "How many years of professional design experience do you have?"
3. "What is your current role?" (with options for various design and decision-making positions)

Only participants meeting inclusion criteria will proceed to the main study.

Will incentives (compensation) be offered for research participation?

Yes

Describe the incentive strategy.

Participants will receive $8-10 for completion of the 30-minute study, equivalent to approximately $15-20 per hour. Payment will be made through the recruitment platform (Prolific) or via digital gift cards for other recruitment channels.

Describe any alternative procedures available to those who choose not to participate.

Participation is entirely voluntary with no alternative procedures needed.

Is it likely that a member of the research team will have an authority relationship with potential participants?

No

Does your research involve recruitment of participants in any of the following countries/regions: United Kingdom, countries in the European Economic Area, China, Taiwan, or South Korea?

Potentially yes - we may recruit internationally through online platforms.

Clinical Trials

Are participants prospectively assigned to an intervention in this study?: Yes (participants are randomly assigned to view either standard or enhanced metrics; all participants see both AI and human evaluation data)

Is the study designed to evaluate the effect of the intervention(s) on participants?: Yes

Is the effect being evaluated a health-related biomedical or behavioral outcome?: No

---

Risks and Benefits

Indicate all anticipated risks of harm/discomfort to participants or others.

Minimal risks:

- Possible mild discomfort when evaluating potentially manipulative interface designs
- Extended time commitment (35-40 minutes)
- Potential professional reflection on ethical design practices and AI tool usage
- Possible cognitive fatigue from evaluating multiple data sources per interface

Describe any potential benefits to participants and/or society in general.

Participant benefits:

- Opportunity to reflect on ethical design practices
- Contribution to research on ethical design
- Modest financial compensation

Societal benefits:

- Advanced understanding of how AI evaluation tools affect design decision-making

- Evidence on whether enhanced ethical metrics work across different evaluation sources
- Insights into designer trust and adoption of AI evaluation tools
- Potential improvements in both AI-assisted and human-centered design evaluation practices

---

Privacy & Confidentiality

Describe how you will protect the privacy of participants throughout the course of the research procedures.

Participants will complete the study in their own private environment. No personally identifiable information will be collected beyond basic demographic and professional information. All data will be stored securely on encrypted Cornell systems.

Will any personally identifiable information (PII) be collected or obtained from or about participants?

No

Will any demographic information be collected which could lead to a deductive disclosure of participant(s) identities?

No - only general demographic information (years of experience, general role type, industry) will be collected.

In what format(s) will research data be created or documented?

Research data will be collected as survey responses in Qualtrics and exported as spreadsheets for analysis. Responses will include quantitative ratings and qualitative text responses.

Describe how research data will be shared among research team members.

Data will be stored on secure Cornell Box drives accessible only to the research team. All data will be de-identified before analysis.

Where will the data be stored?

Data will be stored on Cornell Qualtrics during collection and transferred to secure Cornell Box drives for long-term storage and analysis.

In what format(s) will the data be maintained during the life of the study?

Data will be maintained in spreadsheet format (.csv/.xlsx) on encrypted Cornell systems.

Do you have plans to store samples or study data for future research use?

No

Describe any other aspects of your data management plan.

No additional identifying information will be collected. All analysis will be conducted on aggregated data. Individual responses will not be traceable to specific participants.

Are you aware of any likely ethical or legal circumstances when it would be necessary to break confidentiality?

No

Describe any plans to de-identify and share the research data.

Aggregated, de-identified findings will be shared through academic publications and presentations. No individual-level data will be shared.

Indicate all intended forms of dissemination of research results.

- Journal article
- Academic conference presentations
- Thesis/dissertation
- Potential industry reports or white papers

Will any protected health information (PHI) be collected or obtained?

No

Connected Projects

Indicate if any part of your project is funded by an external sponsor.

No current external funding

Additional Documents

[Survey instrument, recruitment materials, consent form, and interface examples to be attached]