Survey Instrument Outline

Section 1: Consent and Eligibility (2 minutes)

Informed Consent

- Display full consent form
- Required selection: "I consent to participate" OR "I do not consent to participate"
- Non-consenting participants exit study

Eligibility Screening

1. **Professional Experience Check**
    - "Do you have professional experience in UI/UX design, product design, or design decision-making roles?"
    - Options: Yes / No
    - "No" responses exit study
2. **Experience Level**
    - "How many years of professional design experience do you have?"
    - Options: Less than 1 year / 1-2 years / 3-5 years / 6-10 years / More than 10 years
    - "Less than 1 year" responses exit study
3. **Age Verification**
    - "Are you 18 years of age or older?"
    - Options: Yes / No
    - "No" responses exit study

Section 2: Demographics and Professional Background (3 minutes)

Current Role

- "What is your current primary role?"
- Options: UX Designer / UX Researcher / UI Designer / Product Designer / Design Manager / Product Manager / Design Director / Other (specify)

Industry Experience

- "Which industries have you worked in? (Select all that apply)"
- Options: Technology/Software / E-commerce / Healthcare / Financial Services / Gaming / Media/Entertainment / Consulting / Other (specify)

Company Size

- "What is the approximate size of your current organization?"

- Options: Startup (<50 employees) / Small (50-200) / Medium (200-1000) / Large (1000-5000) / Enterprise (5000+) / Freelance/Consultant

Decision-Making Authority

- "In your current role, do you have the authority to make final decisions about interface designs?"
- Options: Yes, final decision authority / Yes, significant influence / Some input / Little input / No decision authority

AI Evaluation Experience

Section 3: Interface Evaluation Task (25-30 minutes)

Instructions

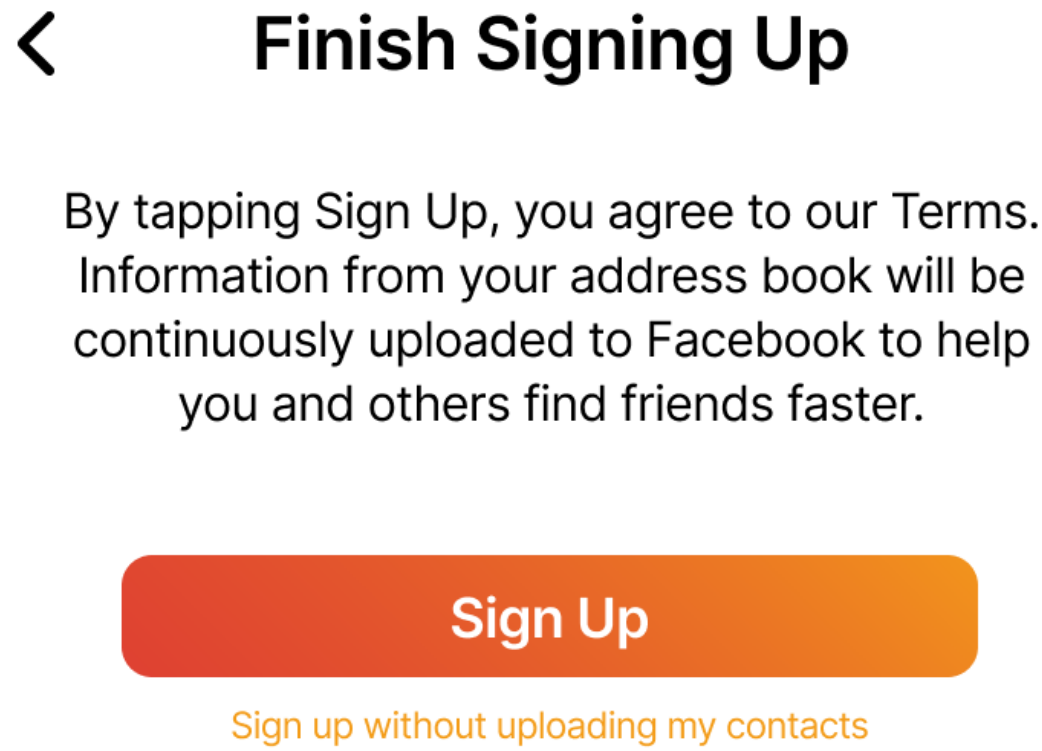
"You will now see a series of interface designs along with user evaluation data. Each interface has been evaluated for user experience. Please review each design and its evaluation data, then answer the questions below. Imagine you are in a position to make or influence the decision about whether to implement this design."

For Each Interface (10-12 interfaces total):

Interface Presentation

- Display interface mockup

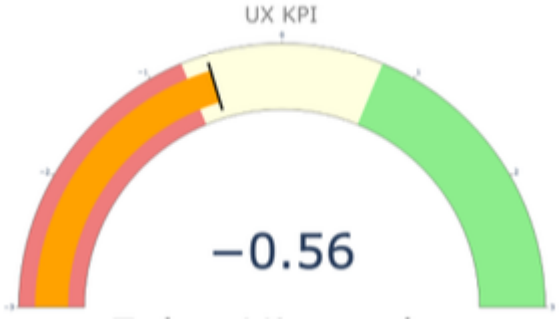
    - Example Mockup Interface:



- Show ONE evaluation dataset per interface:
    - **Half of interfaces**: "Based on evaluation by 20 human participants"
    - **Half of interfaces**: "Based on multiple AI analyses conducted using three different vision-language models trained on UX data, calibrated to approximate the predictive accuracy of evaluations by 20 human participants."
    - Source assignment randomized per participant

Evaluation Data Format (varies by condition)

- **Condition A (Standard UX Metrics)**:
    - Shows: efficiency, satisfaction, ease of use, clarity
        - Example Gauge:



- **Condition B (Enhanced Ethics Metrics)**:
    - Shows: efficiency, satisfaction, ease of use, clarity PLUS manipulation, deception, addiction potential, pressuring behavior

- Example Gauge: 

Primary Decision Questions

- **"How likely would you be to release this interface design based on this evaluation data?"**
- Scale: 1 (Definitely would not release) to 7 (Definitely would release)
- **"Would you release this interface design?"**
- Options: Yes / No

Reasoning

- **"Please explain your decision:"**
- Open text field (300 character limit)

Decision Confidence

- **"How confident are you in the decision you just made?"**
- Scale: 1 (Not at all confident) to 7 (Extremely confident)

Interface Types Include:

1. **Dark Pattern Examples** (from literature):
    - Manipulative subscription flows
    - Deceptive privacy settings
    - Forced social sharing
    - Misleading button labels
2. **Neutral/Positive Examples**:
    - Standard e-commerce checkout
    - Clear navigation designs
    - Transparent privacy controls
3. **Ambiguous Cases**:
    - Gamification elements
    - Social proof indicators
    - Personalization features

Section 4: Decision-Making Process Questions (7 minutes)

Evaluation Source Preferences

- **"Overall, which type of evaluation data did you find more useful for decision-making?"**
- Options: Human evaluation data / AI evaluation data / Both were equally useful / Neither was particularly useful
- **"Which type of evaluation data did you trust more?"**
- Options: Human evaluation data / AI evaluation data / Both equally / Neither

Metric Importance

- **"Which factors were most important in your decision-making? (Rank top 3)"**
- Options: User satisfaction / Ease of use / Business impact / Ethical considerations / Legal compliance / User safety / Innovation / Evaluation source (human vs AI) / Other (specify)

AI vs Human Evaluation Perceptions

- **"How accurate do you believe AI evaluation of interfaces is compared to human evaluation?"**
- Scale: 1 (Much less accurate) to 7 (Much more accurate), with 4 being "About the same"
- **"How reliable do you believe AI evaluation of interfaces is compared to human evaluation?"**
- Scale: 1 (Much less reliable) to 7 (Much more reliable), with 4 being "About the same"

Professional Practice

- **"In your professional work, how often do you encounter designs that you consider ethically questionable?"**
- Options: Never / Rarely / Sometimes / Often / Very often
- **"How familiar are you with the concept of 'dark patterns' in interface design?"**
- Options: Very familiar / Somewhat familiar / Slightly familiar / Not familiar

Organizational Support

- **"How supportive is your organization of using AI tools for design evaluation?"**
- Scale: 1 (Not at all supportive) to 7 (Extremely supportive) + "Not applicable"
- **"How supportive is your organization of rejecting designs based on ethical concerns?"**
- Scale: 1 (Not at all supportive) to 7 (Extremely supportive) + "Not applicable"

Future Tool Preferences

- **"What additional information would be most helpful when evaluating interface designs?"**
- Multiple select: Long-term user behavior data / Accessibility metrics / Privacy impact assessments / Psychological impact measures / Regulatory compliance scores / Combined AI-human evaluation /

Other (specify)

- **"Have you previously used AI tools to evaluate or analyze user interfaces?"**
- Options: Yes, frequently / Yes, occasionally / Yes, but rarely / No, never

Open Feedback

- **"Any additional thoughts about comparing AI vs human evaluation data, or this study in general?"**
- Open text field (optional, 500 character limit)

Section 5: Debrief and Completion (2 minutes)

Study Purpose Revelation

"Thank you for your participation! This study examined two factors: (1) whether presenting user experience data with enhanced ethical metrics influences designers' willingness to implement potentially problematic interfaces, and (2) how AI-generated vs human-generated evaluation data affects design decision-making."

Final Questions

Awareness Check

- **"Did you notice differences in how the user evaluation data was presented across different interfaces?"**
- Options: Yes / No / Unsure
- **"Did you notice patterns in differences between AI and human evaluation data?"**
- Options: Yes / No / Unsure

Study Feedback

- **"Do you have any feedback about this study or questions about the research?"**
- Open text field (optional)

Completion

- Display completion code for compensation
- Provide researcher contact information: Hauke Sandhaus (hgs52@cornell.edu), Faculty Advisor: Helen Nissenbaum (hn288@cornell.edu)
- Thank participant

Technical Implementation Notes

Randomization

- Participants randomly assigned to **Condition A (Standard)** or **Condition B (Enhanced Ethics)** - Between-subjects
- Each participant sees all dark pattern interfaces
- **Data source assignment**: Half interfaces show Human data, half show AI data - Within-subjects
- Interface-to-data-source pairing randomized per participant
- Interface presentation order randomized

Data Collection

- Track which interfaces received Human vs AI data per participant
- Record condition assignment (Standard vs Enhanced Ethics)
- All responses timestamped
- Interface presentation order tracked