CSE 593

Assignment 1 post-mortem
Informed consent
Assignment 2 review

Farnaz Jahanbakhsh

ELECTRICAL ENGINEERING
AND COMPUTER SCIENCE
UNIVERSITY OF MICHIGAN

Logistics

- Working on Assignment 1 (Individual) grades.

- Assignment 1 (Group) was due yesterday at 5PM.

- Assignment 2 assigned today.

- Assignment 2 (Individual) due in two weeks (10/09 at 5PM).

- Assignment 2 (Group) due in three weeks (10/16 at 5PM)

User-centered design process
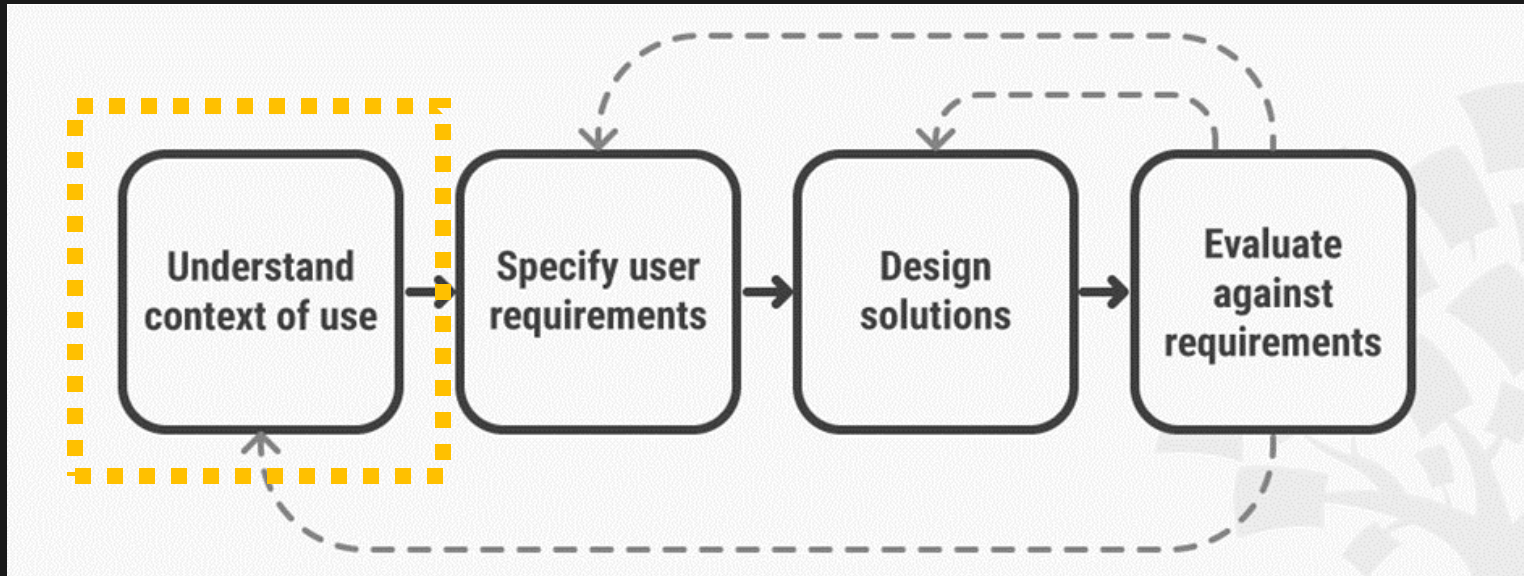
Assignment 1 (individual) post mortem

Think-aloud protocol
- Watch, listen, and take notes
- Don't intervene to "help" the participant
  - Unless they get stuck

- Probe, Ask follow-up questions

Assignment 1 (group) post mortem

- You conducted your surveys! Great job!

- We loved the diverse set of foci.

- Good number of responses.

- Some insightful findings.

Assignment 1 (group) post mortem

- Challenges formulating promise, obstacles, and solution.

- Disconnect between related work and the rest of the assignment. Not clear how it informs understanding of context of use.

Assignment 1 (group) post mortem

- Survey design challenges:
  - No clear inclusion/exclusion criteria
  - No clear informed consent.
  - Insufficient consideration for demographics data.
  - Few questions directly explore breakdowns.
  - Some hypothetical questions.
  - Arbitrary choice of Likert Scales.
  - Analysis rarely explores relationship between answers to different questions.
  - Some challenges with visualizing data.
  - No explicit changes to focus in the discussion.

Informed Consent
vs
Deception

Biases in ratings of online gig workers

Differential performance evaluation

ARE EMILY AND GREG MORE EMPLOYABLE
THAN LAKISHA AND JAMAL?
A FIELD EXPERIMENT ON LABOR MARKET DISCRIMINATION

Marianne Bertrand
Sendhil Mullainathan

What's in a Name: Exposing Gender Bias in Student Ratings of Teaching

Lillian MacNell · Adam Driscoll · Andrea N. Hunt

Differential performance evaluation (online edition)

Potential confounding factor:

performance

Bias in Online Freelance Marketplaces: Evidence from TaskRabbit and Fiverr

Anikó Hannák
Northeastern University
ancsaaa@ccs.neu.edu

Claudia Wagner
GESIS Leibniz Institute for the Social Sciences & U. of Koblenz-Landau
claudia.wagner@gesis.org

David Garcia
ETH Zürich, Switzerland
dgarcia@ethz.ch

Alan Mislove
Northeastern University
amislove@ccs.neu.edu

Markus Strohmaier
GESIS Leibniz Institute for the Social Sciences & U. of Koblenz-Landau
markus.strohmaier@gesis.org

Christo Wilson
Northeastern University
cbw@ccs.neu.edu

ABSTRACT
Online freelancing marketplaces have grown quickly in recent years. In theory, these sites offer workers the ability to earn money without the obligations and potential social biases associated with traditional employment frameworks. In this paper, we study whether two prominent online freelance marketplaces—TaskRabbit and Fiverr—are impacted by racial and gender bias. From these two platforms, we collect 13,500 worker profiles and gather information about workers' gender, race, customer reviews, ratings, and positions in search rankings. In both marketplaces, we find evidence of bias: we find that perceived gender and race are significantly correlated with worker evaluations, which could harm the employment opportunities afforded to the workers. We hope that our study fuels more research on the presence and implications of discrimination in online environments.

INTRODUCTION
Online freelance marketplaces such as Upwork, Care.com, and Freelancer have grown quickly in recent years. These sites facilitate additional income for many workers, and even provide a primary income source for a growing minority. In 2014, it was estimated that 25% of the total workforce in the US was involved in some form of freelancing, and this number is predicted to grow to 40% by 2020 [37, 34].

Online freelancing offers two potential benefits to workers. The first, *flexibility*, stems from workers' ability to decide when they want to work, and what kinds of tasks they are willing to perform [33]. Indeed, online freelancing websites provide job opportunities to workers who may be disenfranchised by the rigidity of the traditional labor market, *e.g.*, new parents who can only spend a few hours working on their laptops at night, or people with

11

Controlled experiments to isolate performance

Why deception?

Thebault-Spieker, Jacob, Daniel Kluver, Maximilian A. Klein, Aaron Halfaker, Brent Hecht, Loren Terveen, and Joseph A. Konstan. "Simulation experiments on (the absence of) ratings bias in reputation systems." *Proceedings of the ACM on Human-Computer Interaction* 1, no. CSCW (2017)

No biases manifested

- Intent to deceive ≠ successful deception
- "evaluating work done for someone else may not trigger enough empathy or ownership to show evaluation bias"

Thebault-Spieker, Jacob, Daniel Kluver, Maximilian A. Klein, Aaron Halfaker, Brent Hecht, Loren Terveen, and Joseph A. Konstan. "Simulation experiments on (the absence of) ratings bias in reputation systems." *Proceedings of the ACM on Human-Computer Interaction* 1, no. CSCW (2017)

Simulated workers

Low performing male	Low performing female
High performing male	High performing female

Farnaz Jahanbakhsh, Justin Cranshaw, Scott Counts, Walter S. Lasecki, and Kori Inkpen. "An experimental study of bias in platform worker ratings: the role of performance quality and gender." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-13. 2020.

Farnaz Jahanbakhsh, Justin Cranshaw, Scott Counts, Walter S. Lasecki, and Kori Inkpen. "An experimental study of bias in platform worker ratings: the role of performance quality and gender." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-13. 2020.

15

Farnaz Jahanbakhsh, Justin Cranshaw, Scott Counts, Walter S. Lasecki, and Kori Inkpen. "An experimental study of bias in platform worker ratings: the role of performance quality and gender." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-13. 2020.

Deceptive elements

- Introduced the workers as real
- Labeling was just a set up to expose workers' performance, and a pretext for asking participant to rate
- Performance of the team ostensibly affected participant's pay

Farnaz Jahanbakhsh, Justin Cranshaw, Scott Counts, Walter S. Lasecki, and Kori Inkpen. "An experimental study of bias in platform worker ratings: the role of performance quality and gender." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-13. 2020.

Participant debriefing after deception

- Hypotheses that were tested
- Procedures for deception
- Reason why deception was necessary

User-centered design process

Contextual design

Contextual Design *Evolved* (https://doi.org/10.2200/S00597ED1V01Y201409HCI024)
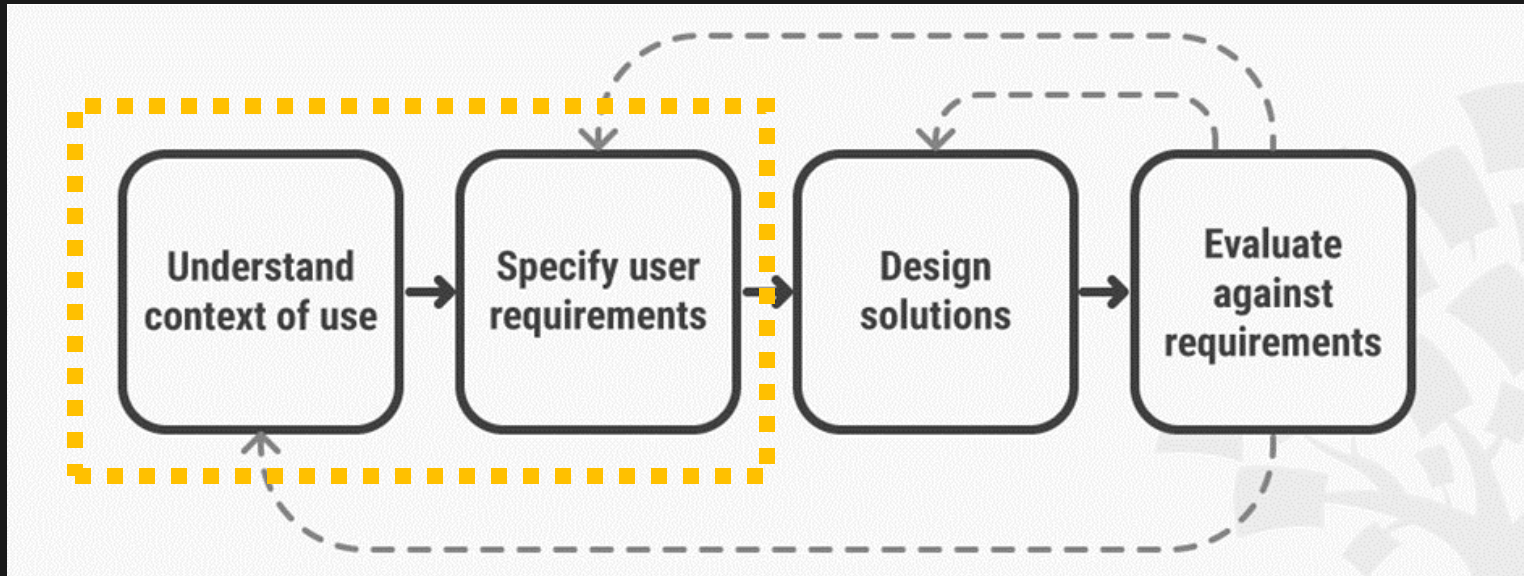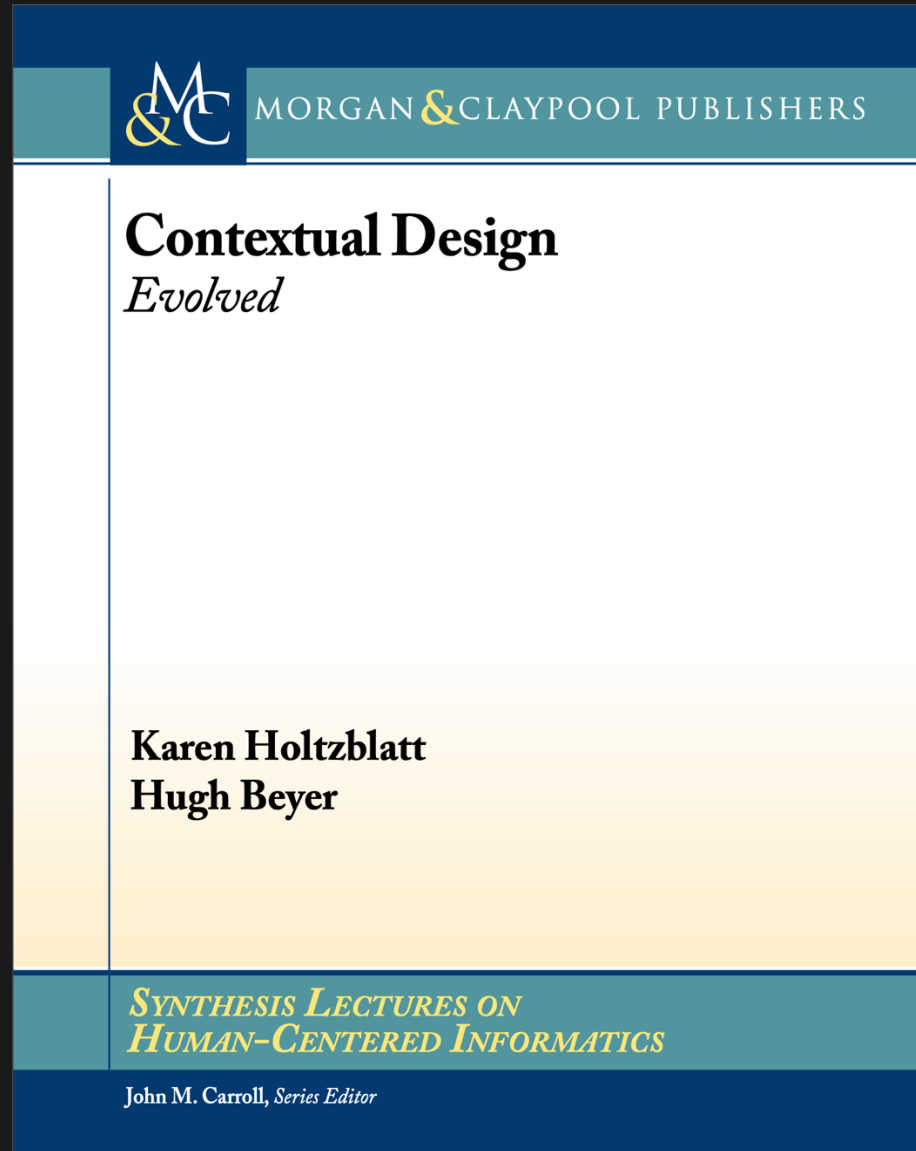
Contextual Design

Contextual Inquiry

Interpretation

Consolidation

Design

Assignment 2 (individual)

- Each member will conduct one Contextual Interview.

- Each member will create interpretations for their interview.

Example: texting a friend

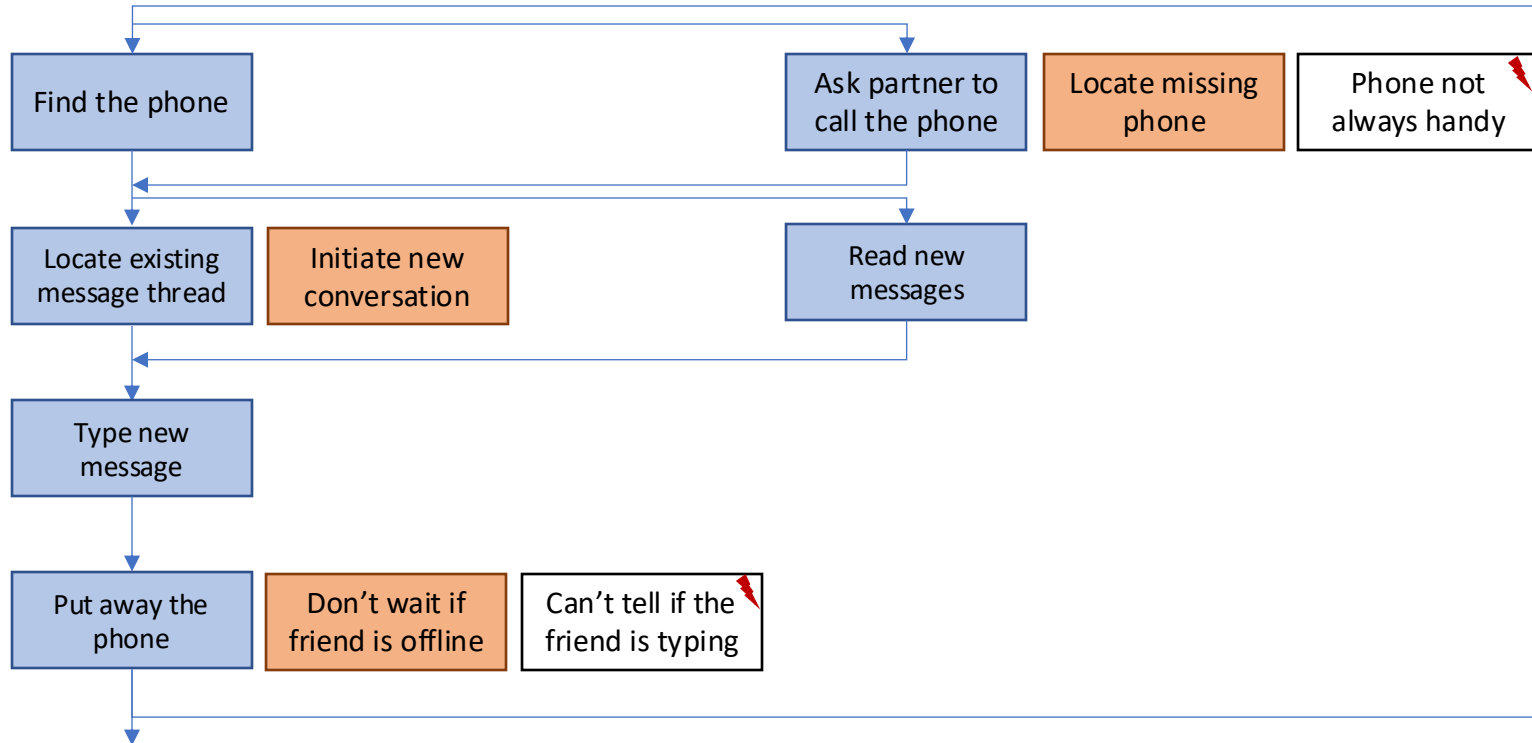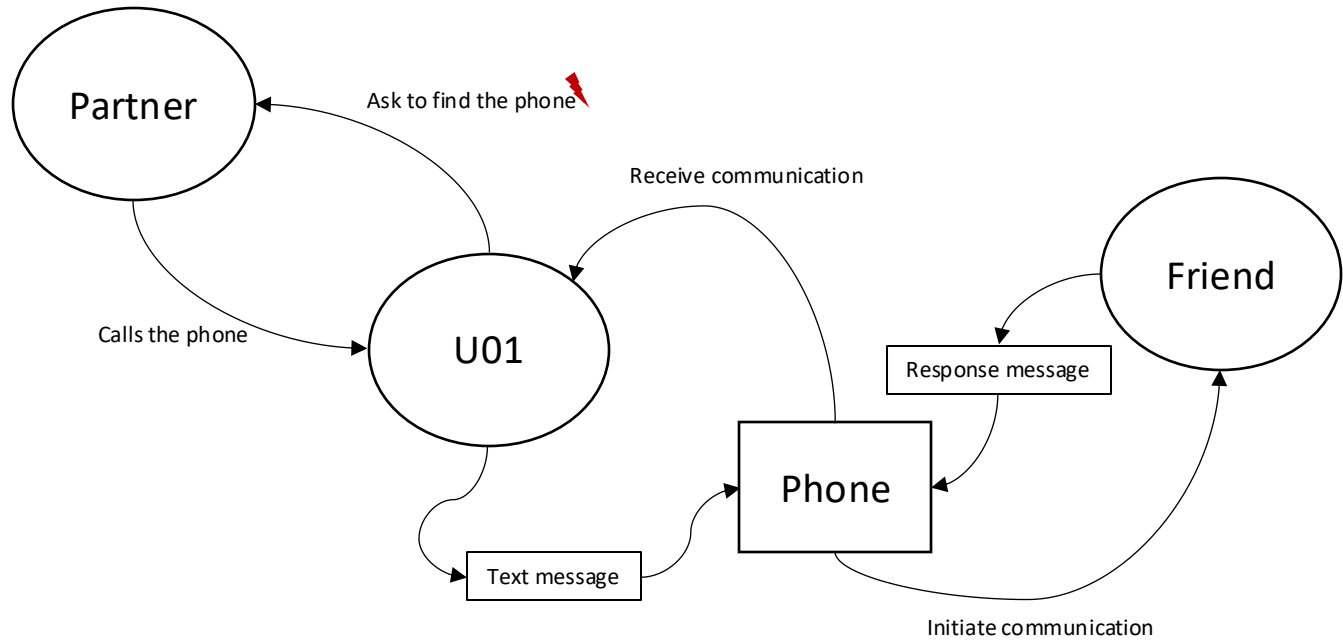| Code	Interpretation
U01-01	They texted their friend to check on existing plans.
U01-02	To text a friend they had to first find their phone that is not always on them.
U01-03	They had to ask their significant other to call their phone to be able to find it.
U01-04	To text a friend they located an old SMS thread and simply "replied" to an unrelated message. It does not matter what the previous message was about.
U01-05	They then typed a message asking about dinner plans later that weekend. They read the message a few time before sending to make sure it was free from errors and that it "made sense."
U01-06	The friend does not always respond right away; that is fine as long as they hear back in time.
U01-07	When phone is handy, they can text right away.

Assignment 2 (individual)

- Each member will conduct one Contextual Interview.

- Each member will create interpretations for their interview.

- Each member will create sequence diagrams and flow diagrams for their participant.

Sequence model

U01: Text a friend to check on existing plans

Flow diagram

Partner

Ask to find the phone

Receive communication

Friend

Calls the phone

U01

Response message

Phone

Text message

Initiate communication

Assignment 2 (group)

- You will work as a team to consolidate your diagrams.

- You will work as a team to create your affinity diagram.

- You will work as a team to come up with requirements.

- But! You may want to start with requirements individually and then consolidate them as a team.

Assignment 2 (group)

- You only have 1 week to do this!

- Keep track of where your other team members are with their individual assignment.

- Encourage everyone to conduct the interview early so that you have the interpretations ready for affinity diagramming.

- Consider starting on the affinity diagram even before you all submit Assignment 2 (Individual)

Questions, comments, and/or concerns?

Farnaz Jahanbakhsh
farnaz@umich.edu
https://people.csail.mit.edu/farnazj