User Interfaces for Fine-grained Integration of Information

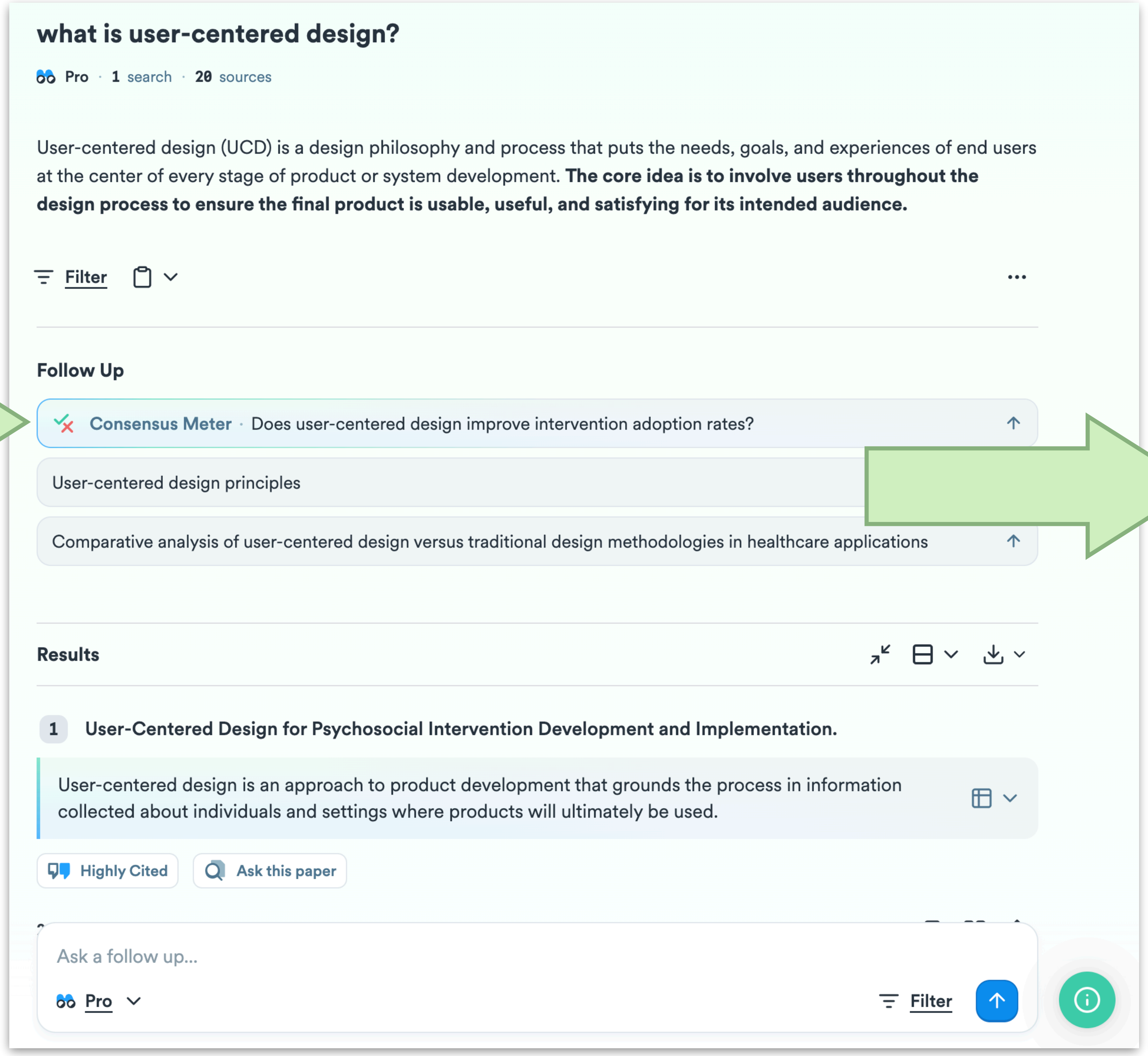
Dissertation Proposal • Monday, August 4, 2025

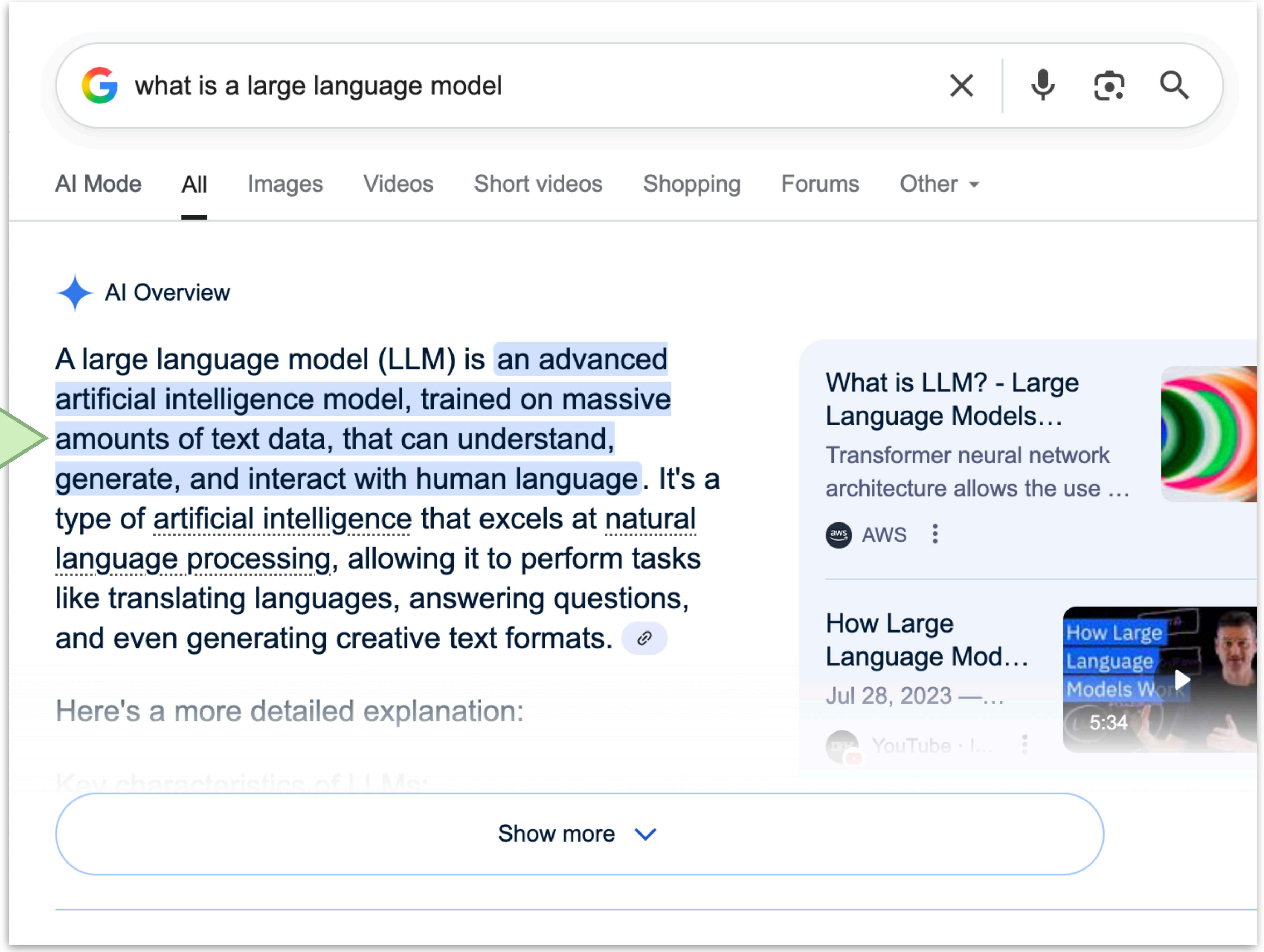
Committee: Anindya De (Chair), Susan Davidson, Insup Lee, Yale Cohen (PSOM)

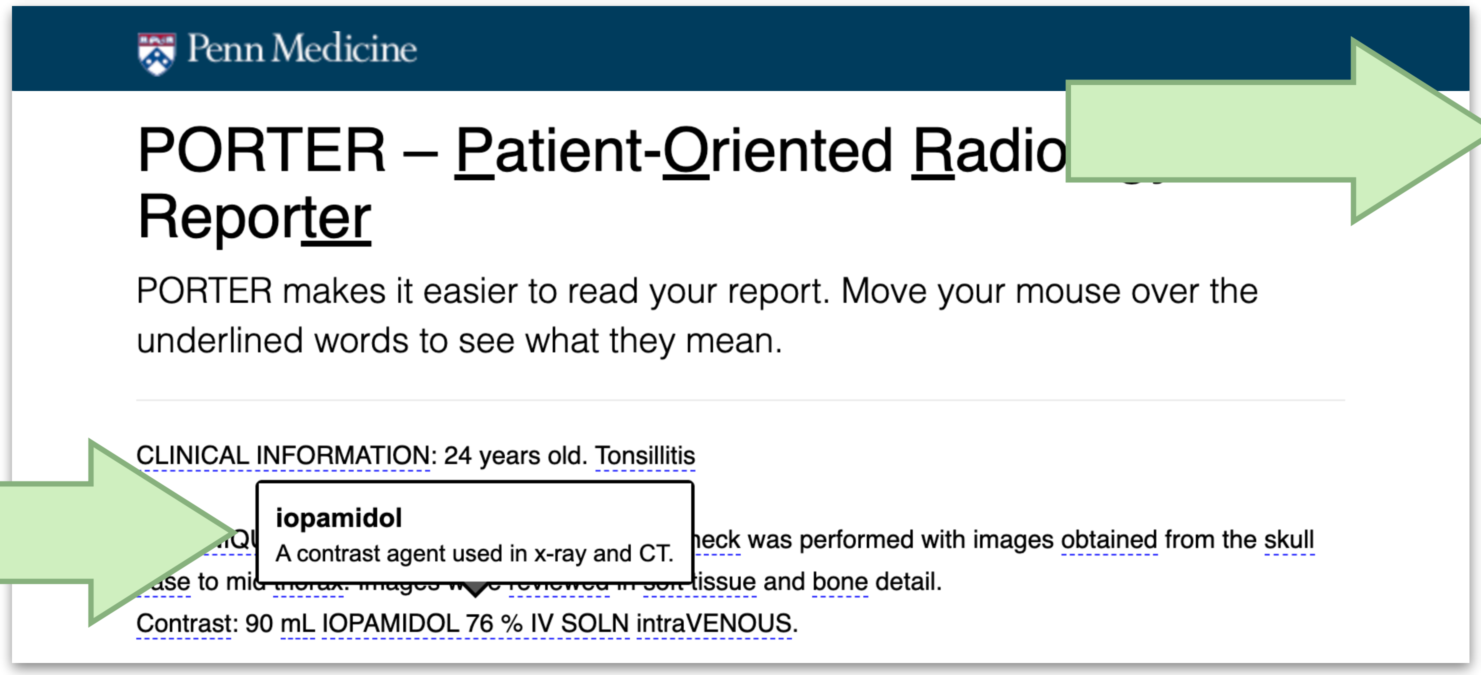
Alyssa Hwang

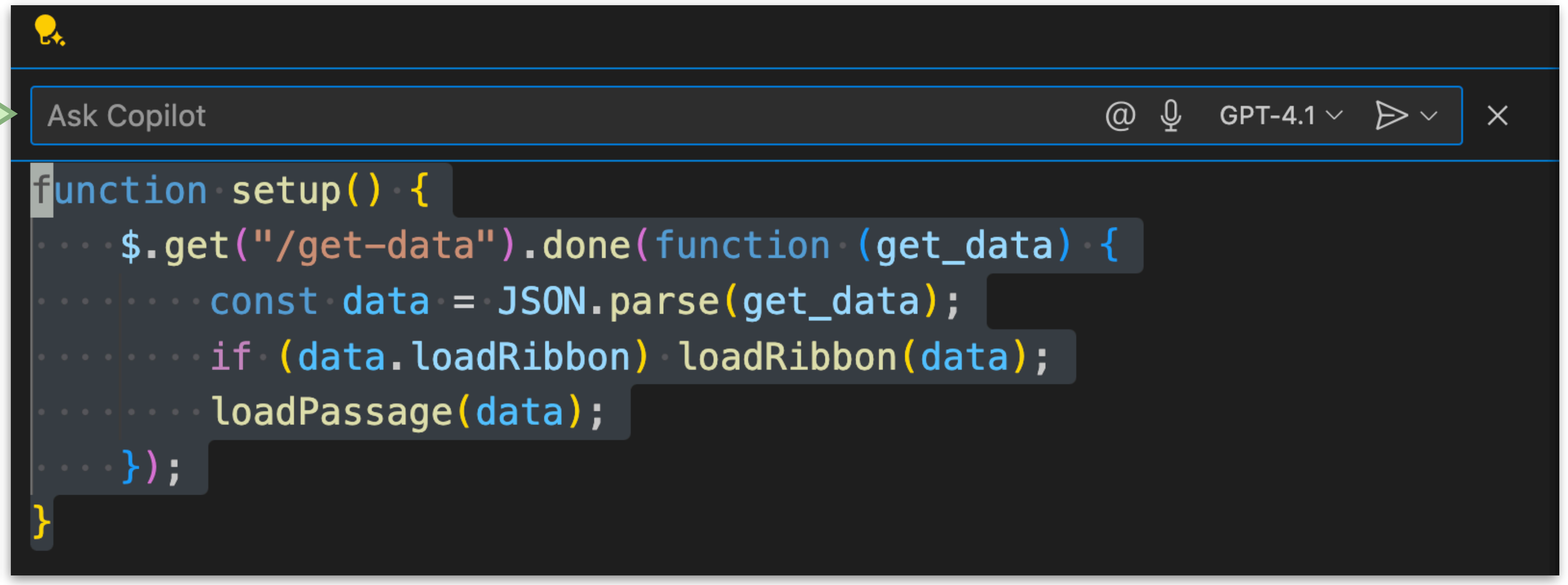
New AI capabilities enable opportunities to enhance information-heavy sources.

Meta
Gemini
deepseek
perplexity
Mistral AI
ANTHROP\C
Copilot
Claude
cohere
stability.ai

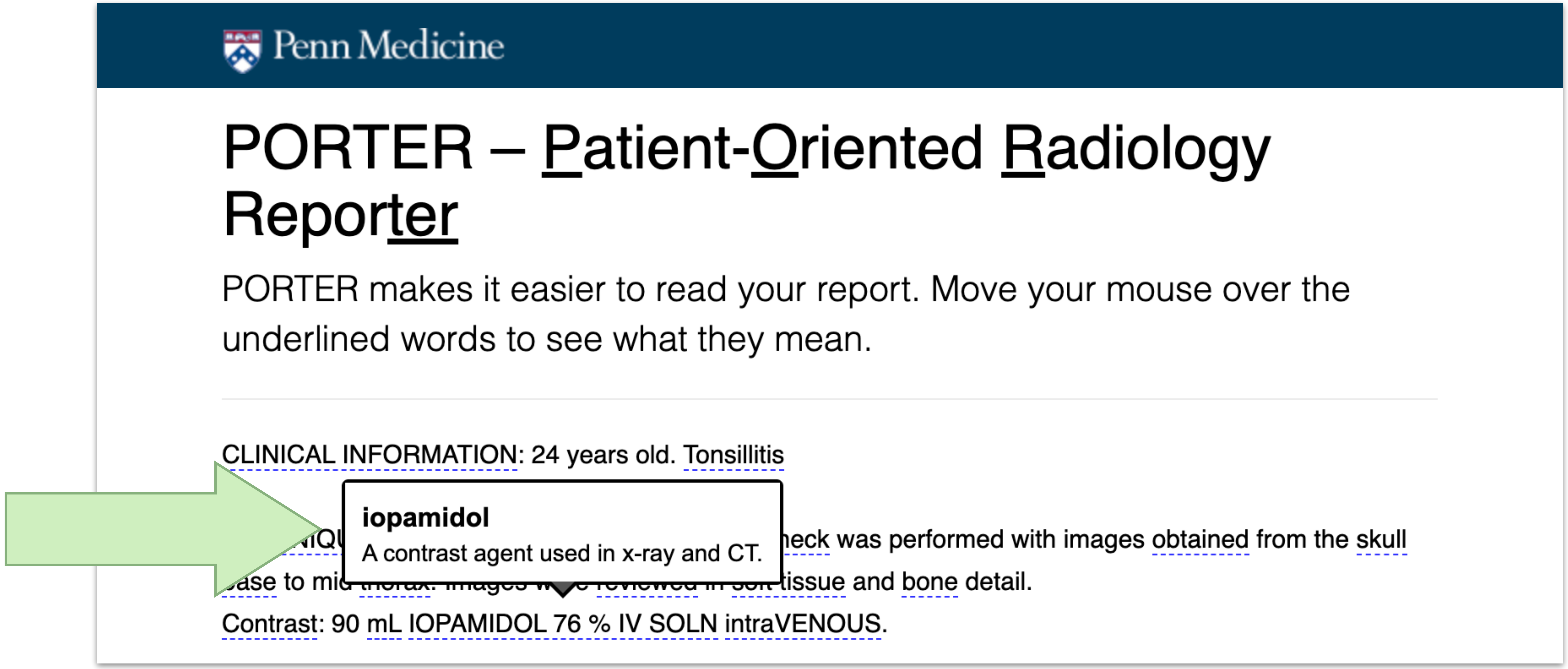








Recent AI advancements support unprecedented levels of work in fine-grained augmentations.



In this thesis, I aim to show that fine-grained augmentations can help users understand information-dense documents.

user-centered approach

Presentation Structure

1. Needs-finding study
2. Case study of prototype
3. Behavioral analysis
4. Proposed work + timeline

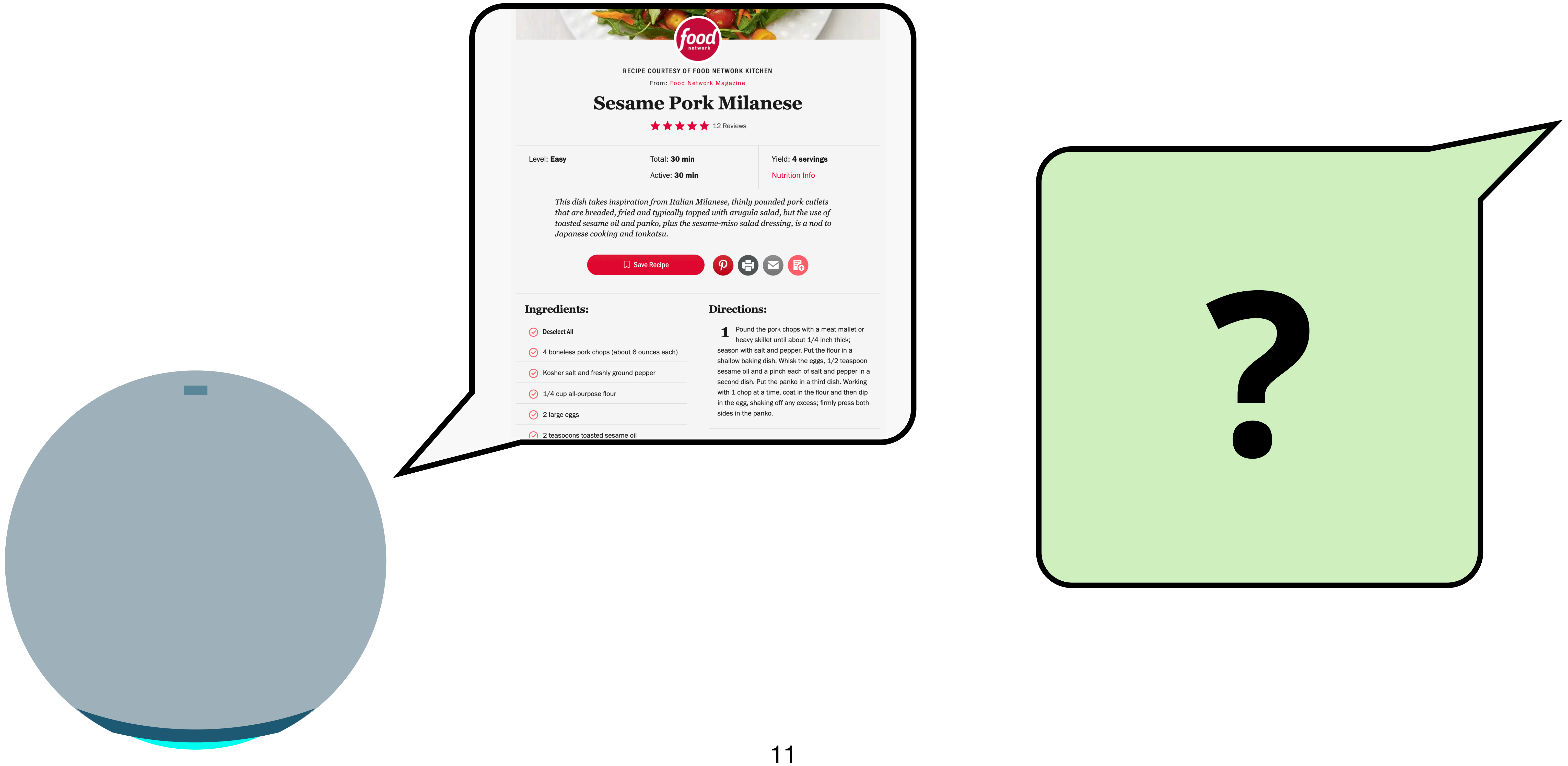
Questions are welcome after each section

1. Needs-finding study

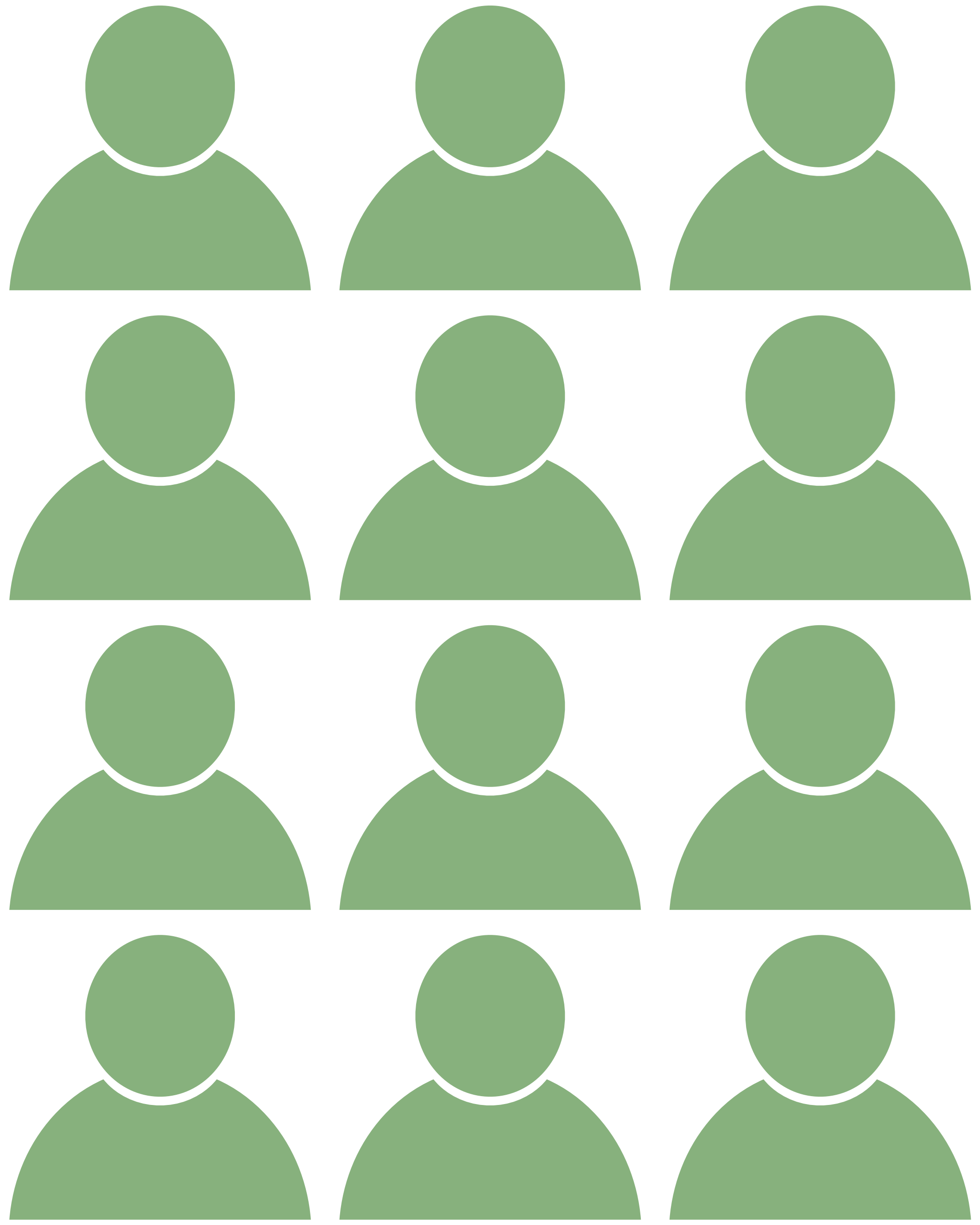
Motivation: understand user needs for fine-grained augmentations

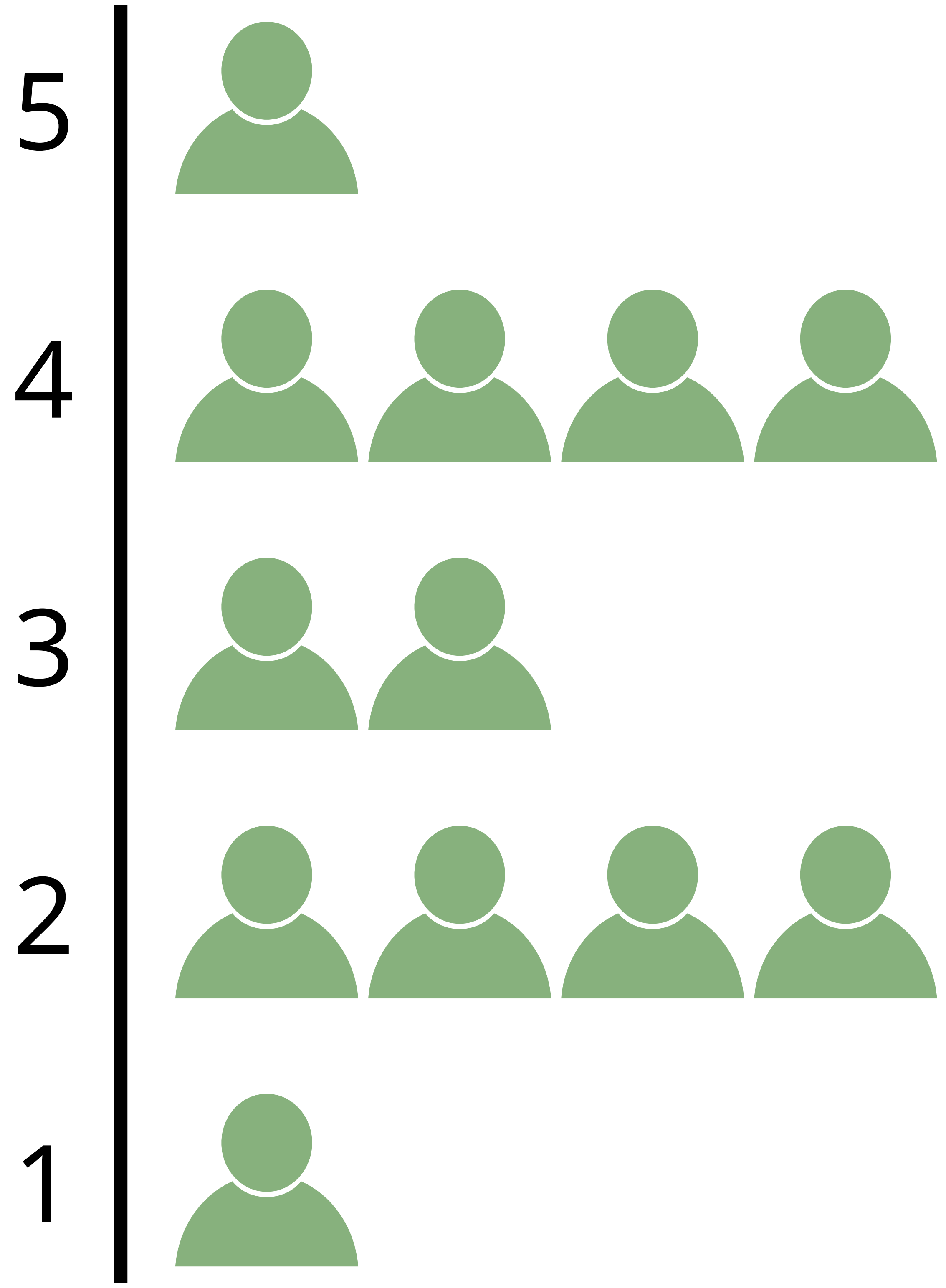
Method: observe users cooking at home with voice assistant

Outcome: definition of user challenges and potential augmentations

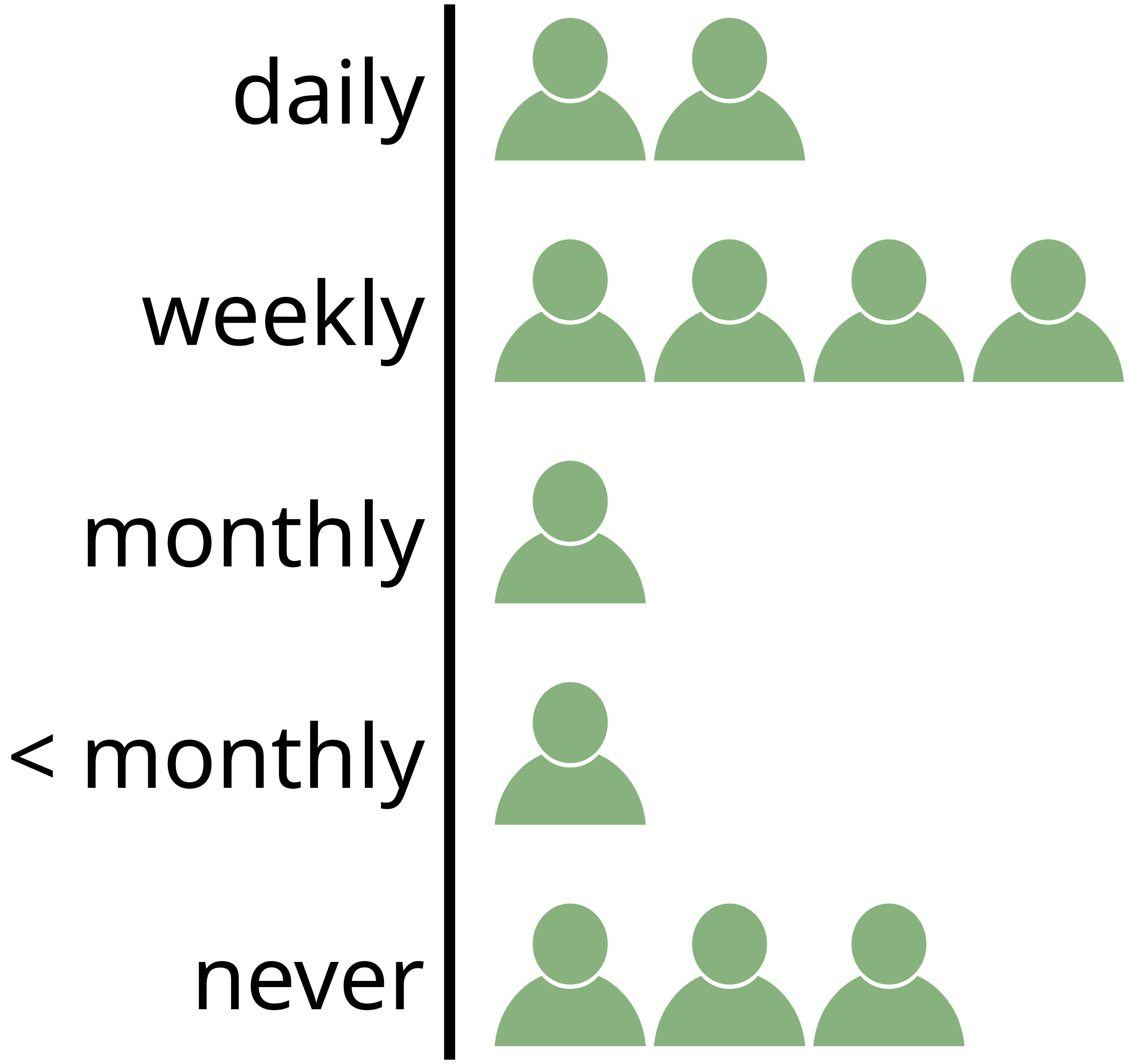


We recruited 12 participants to cook with an Alexa Echo Dot.



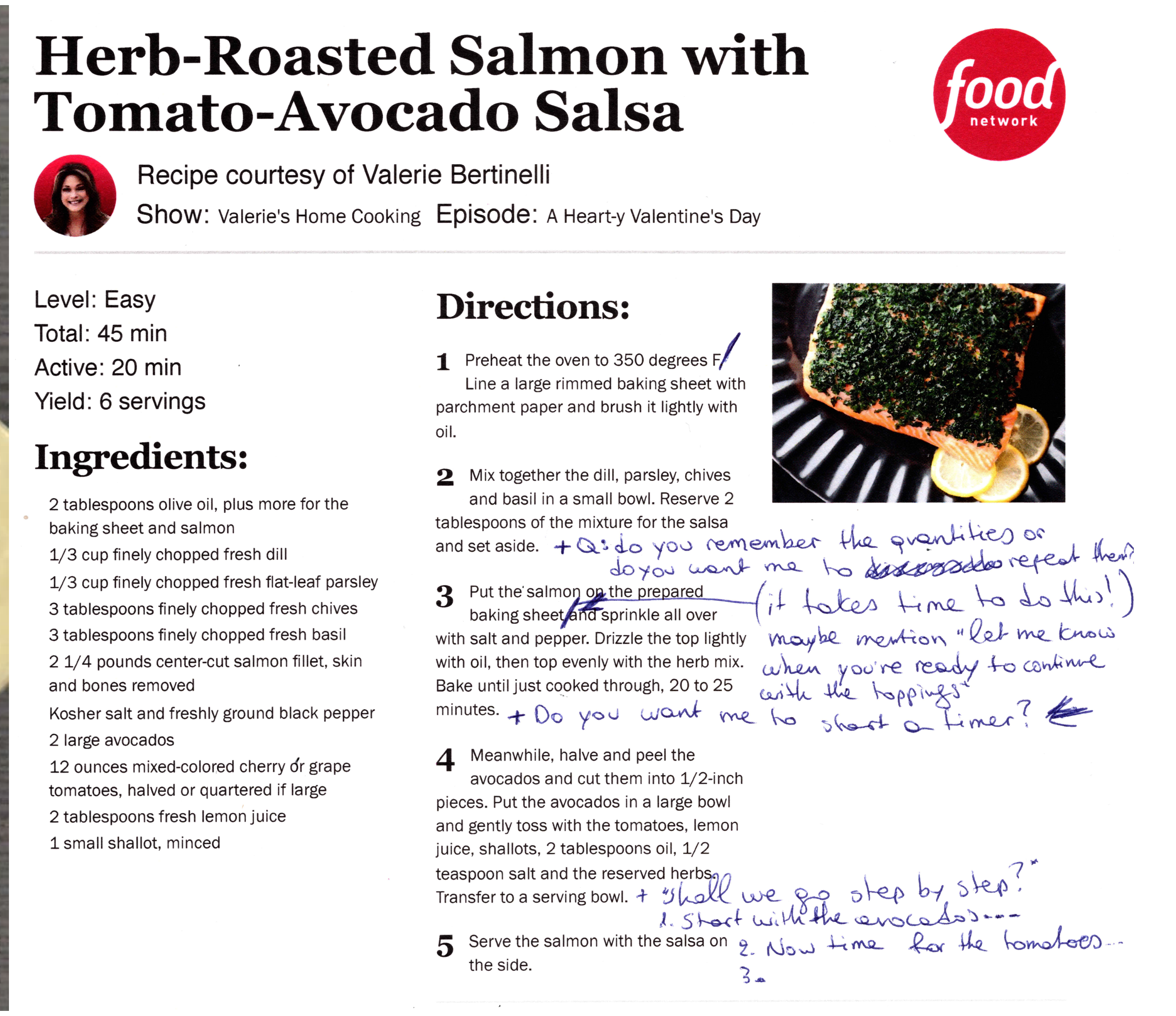


Self-Reported Cooking Skill



Prior Voice Assistant Use

Participants cooked and annotated recipes of their choice in their homes.



Herb-Roasted Salmon with Tomato-Avocado Salsa

Recipe courtesy of Valerie Bertinelli

Show: Valerie's Home Cooking Episode: A Heart-y Valentine's Day

Level: Easy
Total: 45 min
Active: 20 min
Yield: 6 servings

Ingredients:

- 2 tablespoons olive oil, plus more for the baking sheet and salmon
- 1/3 cup finely chopped fresh dill
- 1/3 cup finely chopped fresh flat-leaf parsley
- 3 tablespoons finely chopped fresh chives
- 3 tablespoons finely chopped fresh basil
- 2 1/4 pounds center-cut salmon fillet, skin and bones removed
- Kosher salt and freshly ground black pepper
- 2 large avocados
- 12 ounces mixed-colored cherry or grape tomatoes, halved or quartered if large
- 2 tablespoons fresh lemon juice
- 1 small shallot, minced

Directions:

1 Preheat the oven to 350 degrees F. Line a large rimmed baking sheet with parchment paper and brush it lightly with oil.

2 Mix together the dill, parsley, chives and basil in a small bowl. Reserve 2 tablespoons of the mixture for the salsa and set aside. + Q: do you remember the quantities or do you want me to ~~[illegible]~~ repeat them? (it takes time to do this!)

3 Put the salmon on the prepared baking sheet and sprinkle all over with salt and pepper. Drizzle the top lightly with oil, then top evenly with the herb mix. Bake until just cooked through, 20 to 25 minutes. + Do you want me to start a timer?

maybe mention "let me know when you're ready to continue with the toppings"

4 Meanwhile, halve and peel the avocados and cut them into 1/2-inch pieces. Put the avocados in a large bowl and gently toss with the tomatoes, lemon juice, shallots, 2 tablespoons oil, 1/2 teaspoon salt and the reserved herbs. Transfer to a serving bowl. + "Shall we go step by step?"
1. Start with the avocados...
2. Now time for the tomatoes...
3.

5 Serve the salmon with the salsa on the side.

We observed 9 types of challenges during the user study sessions.

- Missing the big picture
- Information overload
- Fragmentation
- Time insensitivity
- Missing details
- Discarded context
- Failure to listen
- Uncommunicated affordances
- Limitations of audio

We observed 9 types of challenges during the user study sessions.

- Missing the big picture
- Information overload
- Fragmentation
- Time insensitivity
- Missing details
- Discarded context
- Failure to listen
- Uncommunicated affordances
- Limitations of audio

> Put the avocados in a large bowl and gently toss with the tomatoes, lemon juice, shallots, 2 tablespoons oil, 1/2 teaspoon salt and the reserved herbs. Transfer to a serving bowl.

Herb-Roasted Salmon with Tomato-Avocado Salsa, Food Network

We observed 9 types of challenges during the user study sessions.

- Missing the big picture
- Information overload
- Fragmentation
- Time insensitivity
- Missing details
- Discarded context
- Failure to listen
- Uncommunicated affordances
- Limitations of audio

Put the avocados in a large bowl and gently toss with the tomatoes, lemon juice, shallots, 2 tablespoons oil, 1/2 teaspoon salt and the reserved herbs. Transfer to a serving bowl.

Herb-Roasted Salmon with Tomato-Avocado Salsa, Food Network

We observed 9 types of challenges during the user study sessions.

- Missing the big picture
- Information overload
- Fragmentation
- Time insensitivity
- Missing details
- Discarded context
- Failure to listen
- Uncommunicated affordances
- Limitations of audio

Put the avocados in a large bowl and gently toss with the tomatoes, lemon juice, shallots, 2 tablespoons oil, 1/2 teaspoon salt and the reserved herbs. Transfer to a serving bowl.

Herb-Roasted Salmon with Tomato-Avocado Salsa, Food Network

- Missing th
- Informati
- Fragment
- Time inser
- Missing de
- Discarded
- Failure to
- Uncommu
- Limitation

Ingredients:

- Deselect All
- 2 tablespoons olive oil, plus more for the baking sheet and salmon
- 1/3 cup finely chopped fresh dill
- 1/3 cup finely chopped fresh flat-leaf parsley
- 3 tablespoons finely chopped fresh chives
- 3 tablespoons finely chopped fresh basil
- 2 1/4 pounds center-cut salmon fillet, skin and bones removed
- Kosher salt and freshly ground black pepper
- 2 large avocados
- 12 ounces mixed-colored cherry or grape tomatoes, halved or quartered if large
- 2 tablespoons fresh lemon juice
- 1 small shallot, minced

Directions:

1 Preheat the oven to 350 degrees F. Line a large rimmed baking sheet with parchment paper and brush it lightly with oil.

2 Mix together the dill, parsley, chives and basil in a small bowl. Reserve 2 tablespoons of the mixture for the salsa and set aside.

3 Put the salmon on the prepared baking sheet and sprinkle all over with salt and pepper. Drizzle the top lightly with oil, then top evenly with the herb mix. Bake until just cooked through, 20 to 25 minutes.

4 Meanwhile, halve and peel the avocados and cut them into 1/2-inch pieces. Put the avocados in a large bowl and gently toss with the tomatoes, lemon juice, shallots, 2 tablespoons oil, 1/2 teaspoon salt and the reserved herbs. Transfer to a serving bowl.

5 Serve the salmon with the salsa on the side.

ge bowl
the
hallots,
aspoon
herbs.
bowl.

Herb-Roasted Salmon with Tomato-Avocado Salsa, Food Network

We propose 8 augmentations to enhance the source material.

- Missing the big picture
- Information overload
- Fragmentation
- Time insensitivity
- Missing details
- Discarded context
- Failure to listen
- Uncommunicated affordances
- Limitations of audio

- Summarize
- Signpost
- Split
- Elaborate
- Volunteer
- Reorder
- Redistribute
- Visualize

We propose 8 augmentations to enhance the source material.

- Missing the big picture
- Information overload
- Fragmentation
- Time insensitivity
- Missing details
- Discarded context
- Failure to listen
- Uncommunicated affordances
- Limitations of audio

- Summarize
- Signpost
- Split
- Elaborate
- Volunteer
- Reorder
- Redistribute
- Visualize

We propose 8 augmentations to enhance the source material.

- Missing the b
- Information c
- Fragmentatio
- Time insensit
- Missing detai
- Discarded co
- Failure to list
- Uncommunicated affordances
- Limitations of audio

- Visualize

Put the avocados in a large bowl and gently toss with the tomatoes, lemon juice, shallots, 2 tablespoons oil, 1/2 teaspoon salt and the reserved herbs. Transfer to a serving bowl.

Herb-Roasted Salmon with Tomato-Avocado Salsa, Food Network

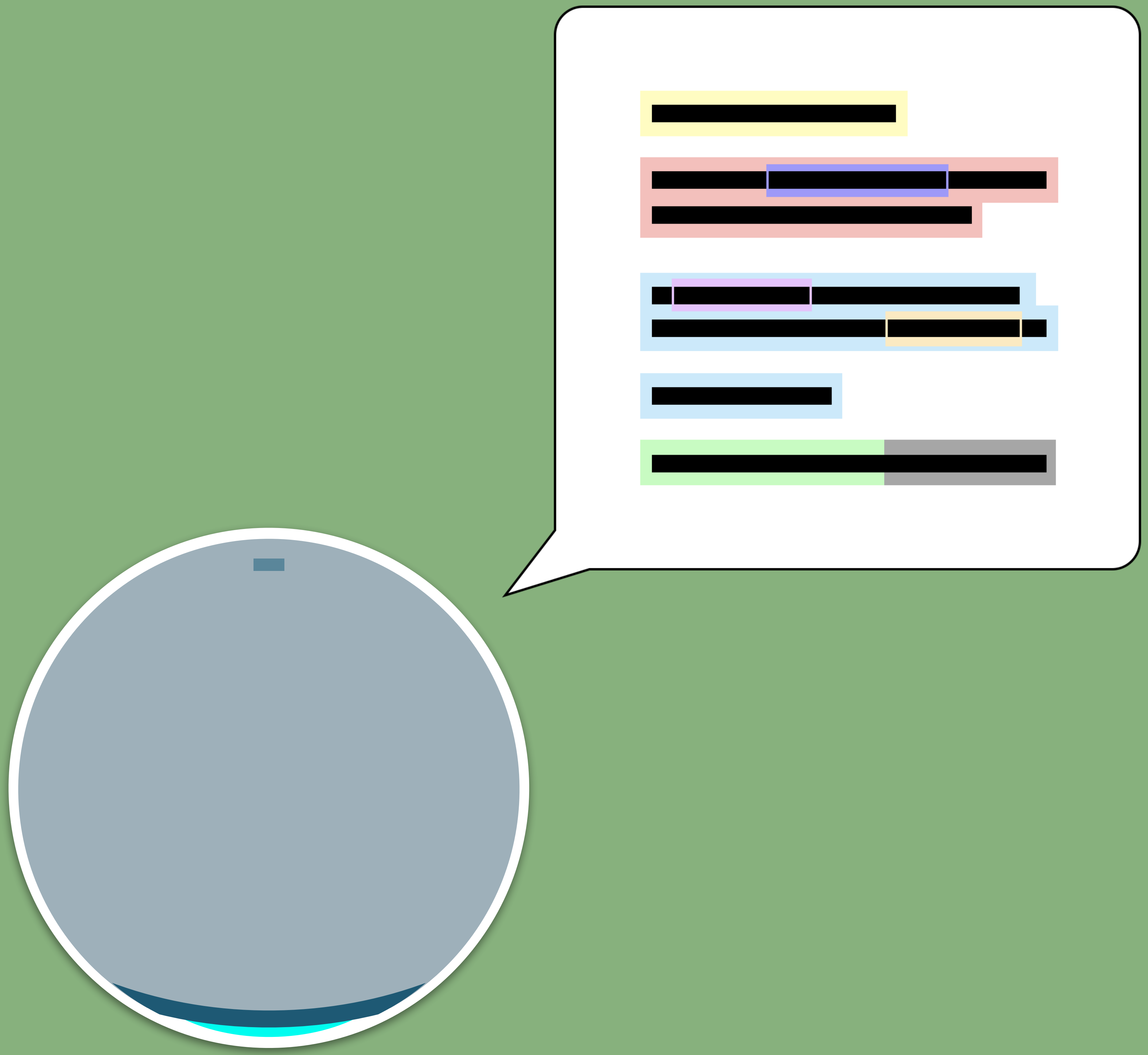
We propose 8 augmentations to enhance the source material.

- Missing the b
- Information c
- Fragmentatio
- Time insensit
- Missing detai
- Discarded co
- Failure to list
- Uncommunic
- Limitations of audio

Put the avocados in a large bowl and gently toss with the 12 ounces of halved tomatoes, 2 tablespoons lemon juice, 1 small diced shallots, 2 tablespoons oil, 1/2 teaspoon salt and the reserved herbs. Transfer to a serving bowl.

Herb-Roasted Salmon with Tomato-Avocado Salsa, Food Network

Needs-finding study outcome



User challenges grounded in realistic study scenario

Potential fine-grained augmentations

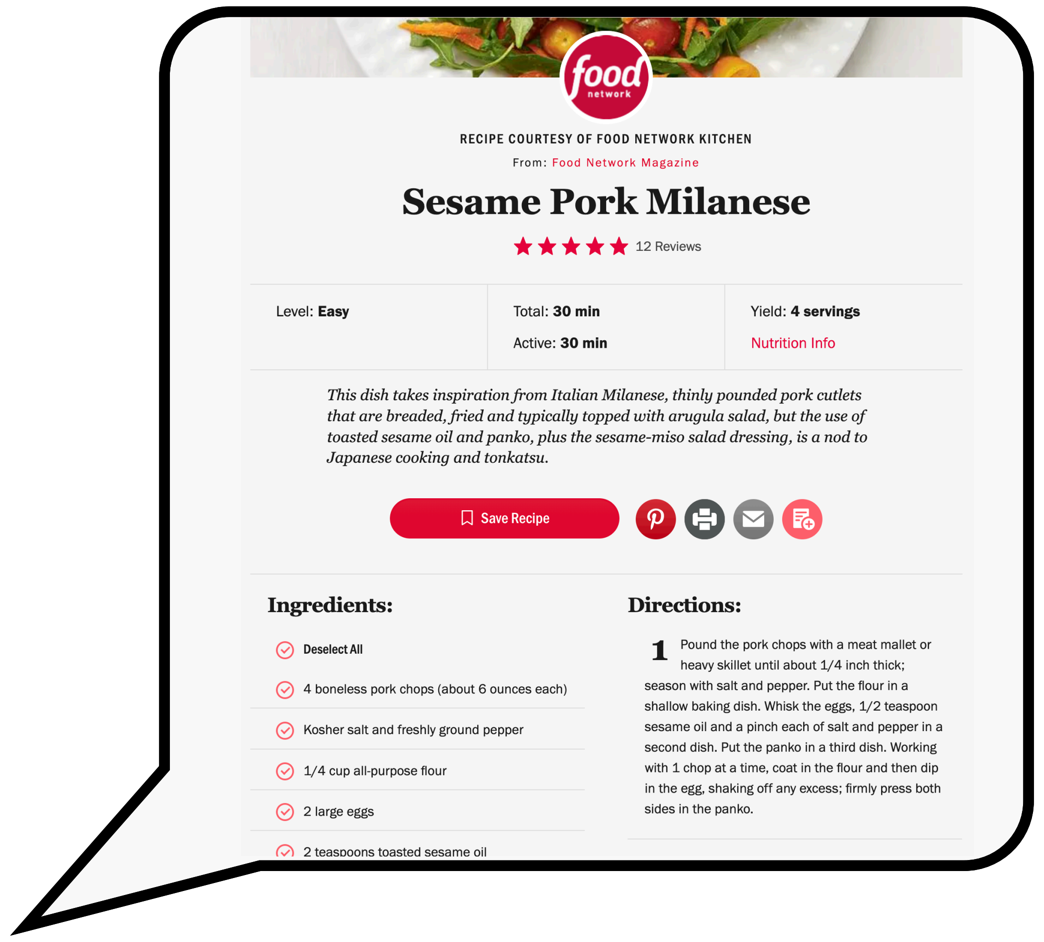
“Rewriting the Script: Adapting Text Instructions for Voice Interaction” (DIS 2023)

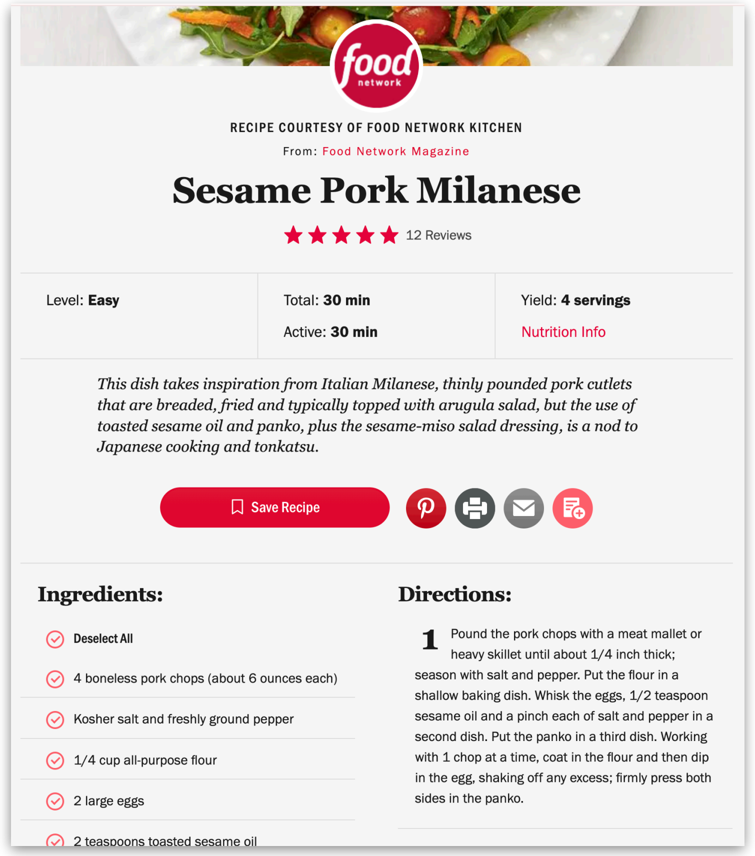
2. Case study of prototype

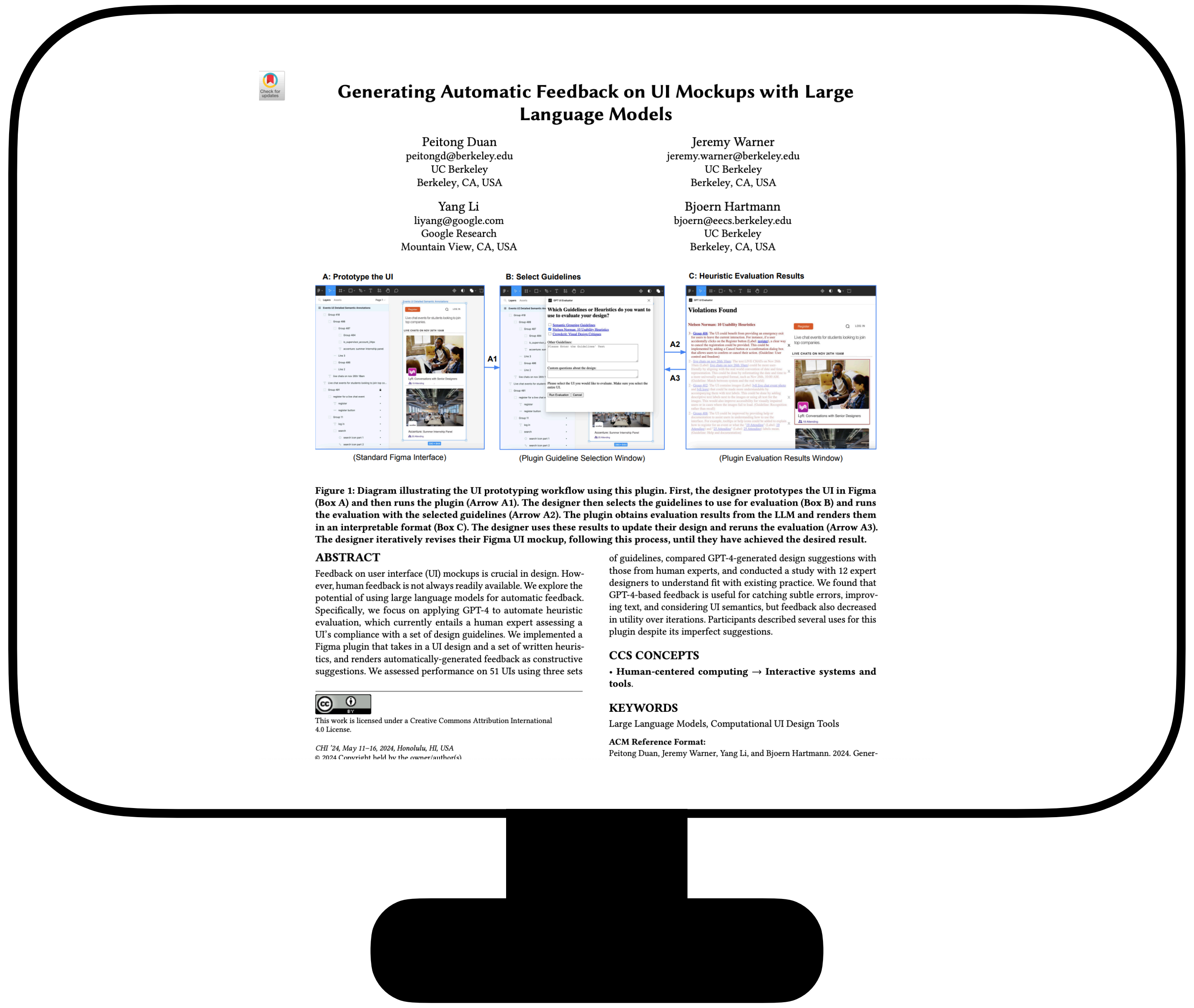
Motivation: observe fine-grained augmentations for real task

Method: implement prototype to support reading info-dense docs

Outcome: augmented reading interface for research papers







Hide Links Clear Selection Reset Zoom

Generating Automatic Feedback on UI Mockups with Large Language Models

Peitong Duan, EECS, UC Berkeley, United States, peitongd@berkeley.edu
Jeremy Warner, EECS, UC Berkeley, United States, jeremy.warner@berkeley.edu
Yang Li, Google Research, United States, yangli@acm.org
Bjoern Hartmann, EECS, UC Berkeley, United States, bjoern@eecs.berkeley.edu

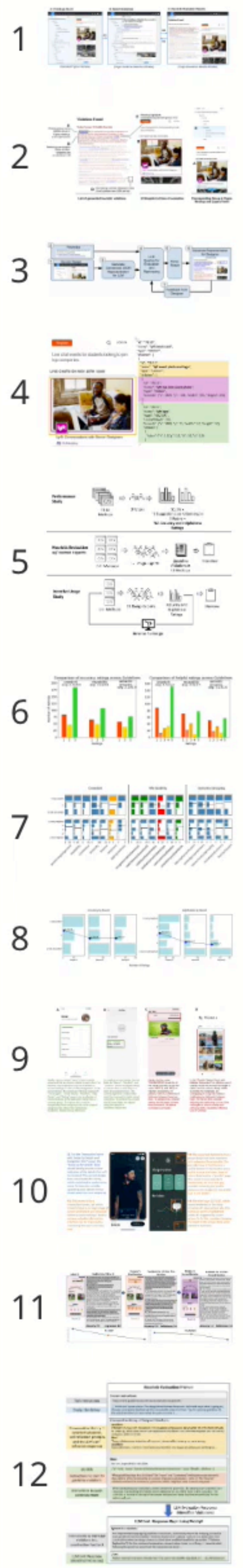
DOI: https://doi-org.proxy.library.upenn.edu/10.1145/3613904.3642782
CHI '24: Proceedings of the CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, May 2024

Feedback on user interface (UI) mockups is crucial in design. However, human feedback is not always readily available. We explore the potential of using large language models for automatic feedback. Specifically, we focus on applying GPT-4 to automate heuristic evaluation, which currently entails a human expert assessing a UI's compliance with a set of design guidelines. We implemented a Figma plugin that takes in a UI design and a set of written heuristics, and renders automatically-generated feedback as constructive suggestions. We assessed performance on 51 UIs using three sets of guidelines, compared GPT-4-generated design suggestions with those from human experts, and conducted a study with 12 expert designers to understand fit with existing practice. We found that GPT-4-based feedback is useful for catching subtle errors, improving text, and considering UI semantics, but feedback also decreased in utility over iterations. Participants described several uses for this plugin despite its imperfect suggestions.

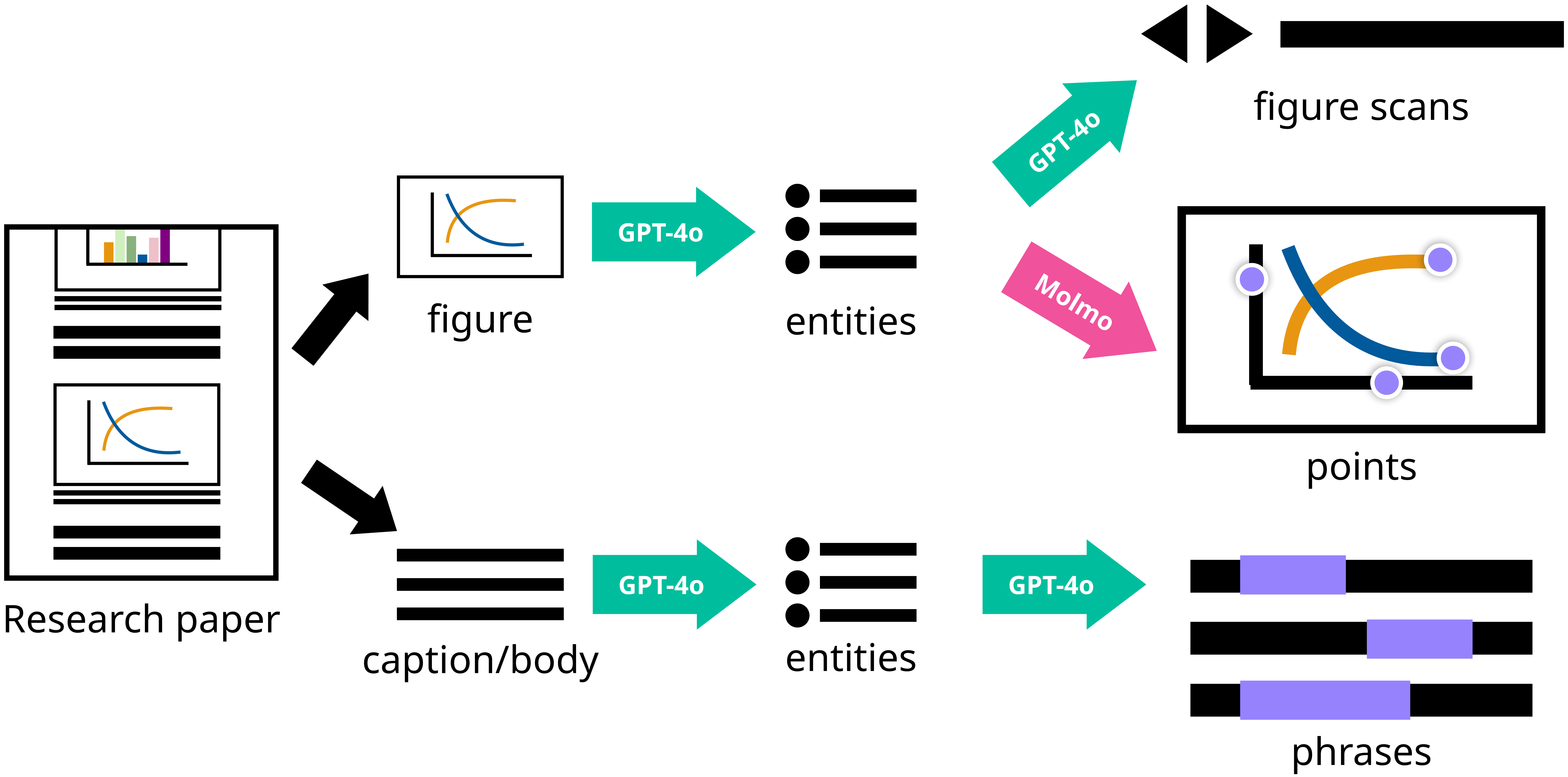
CCS Concepts: • **Human-centered computing** → **Interactive systems and tools**;

Keywords: Large Language Models, Computational UI Design Tools

ACM Reference Format:
Peitong Duan, Jeremy Warner, Yang Li, and Bjoern Hartmann. 2024. Generating Automatic Feedback on UI Mockups with Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11--16, 2024, Honolulu, HI, USA*. ACM, New York, NY, USA 20 Pages. https://doi-org.proxy.library.upenn.edu/10.1145/3613904.3642782



The features in our prototype were generated with a backend AI pipeline.



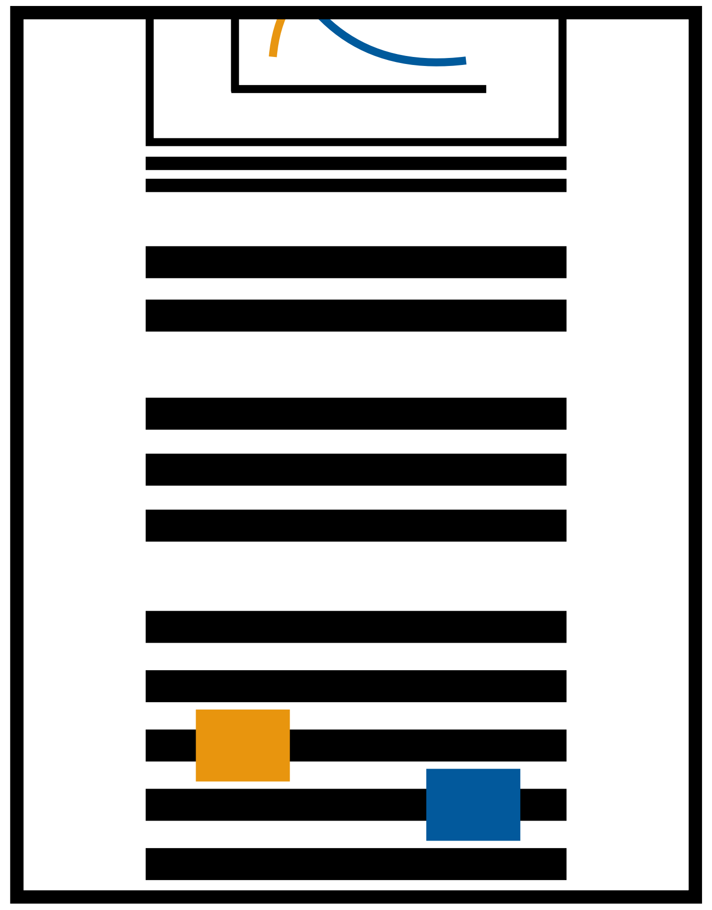
For the formative study, we recruited 10 participants representing a range of fields and reading preferences.

- PhD, physics, laptop + tablet
- PhD, history of science & technology, laptop
- PhD, operations, printed
- PhD, operations, laptop + notepad
- PhD, CS, tablet
- PhD, CS, laptop + notepad
- PhD, CS, laptop
- PhD, CS, laptop
- PhD, CS, printed
- Postdoc, anthropology, printed

We discovered several reading challenges.

Fragmentation Complexity Interpretation Visibility Searching

We discovered several reading challenges.



Visibility

Figure 1: Diagram illustrating the UI prototyping workflow using this plugin. First, the designer p…es the UI in Figma (Box A) and then runs the plugin (Arrow A1). The designer then selects the guidelines to use for evaluation (Box B) and runs the evaluation with the selected guidelines (Arrow A2). The plugin obtains evaluation results from the LLM and renders them in an i…etable format (Box C). The designer uses these results to update their design and reruns …ation (Arrow A3). The designer iteratively revises their Figma UI mockup, following this process, until they have achieved the desired result.

1 …RODUCTION

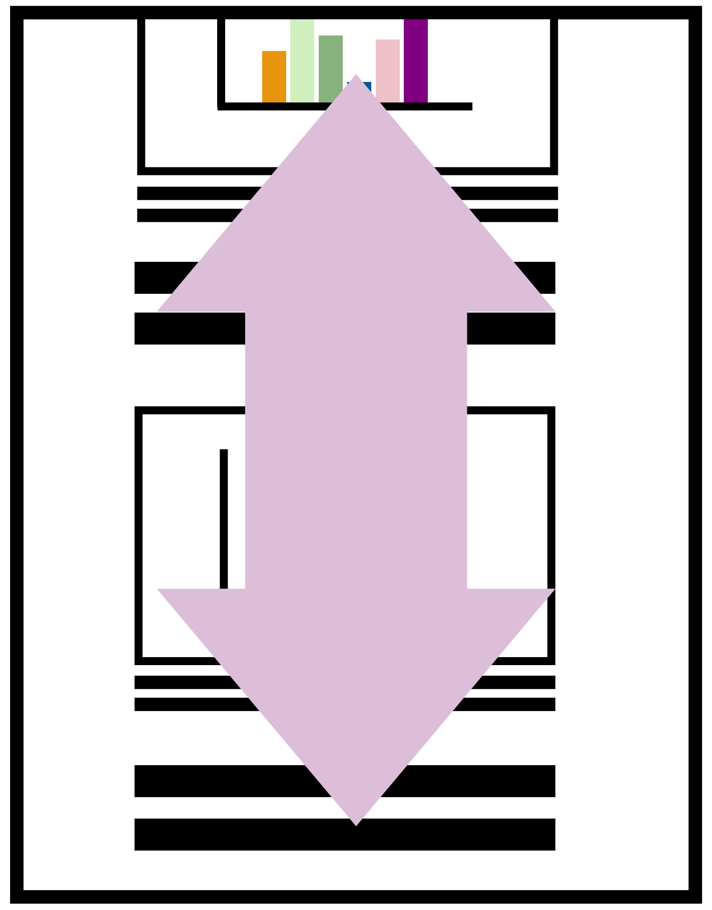
Us… face (UI) design is an essential domain that shapes how humans interact with technology and digital in…on. Designing user interfaces commonly involves iterative rounds of feedback and revision. Feedback is es… or guiding designers towards improving their UIs. While this feedback traditionally comes from humans (v…tudies and expert evaluations), recent advances in computational UI design enable automated feedback. H… automated feedback is often limited in scope (e.g., the metric could only evaluate layout complexity) and ca…llenging to interpret [50]. While human feedback is more informative, it is not readily available and re…me and resources for recruiting and compensating participants.

O… od of evaluation that still relies on human participants today is *heuristic evaluation*, where an ex…ed evaluator checks an interface against a list of usability heuristics (rules of thumb) developed over ti… as Nielsen's 10 Usability Heuristics [39]. Despite appearing straightforward, heuristic evaluation is ch…g and subjective [40], dependent on the evaluator's previous training and personality-related factors [2…e limitations further suggest an opportunity for AI-assisted evaluation.

Th… several reasons why LLMs could be suitable for automating heuristic evaluation. The evaluation process pr… involves rule-based reasoning, which LLMs have shown capacity for [42]. Moreover, design guidelines ar…minately in text form, making them amenable for LLMs, and the language model could also return its fe…s text-based explanations that designers prefer [23]. Finally, LLMs have demonstrated the ability to ur…d and reason with mobile UIs [56], as well as generalize to new tasks and data [28, 49]. However, there ar…asons that suggest caution for using LLMs for this task. For one, LLMs only accept text as input, while us…faces are complex artifacts that combine text, images, and UI components into hierarchical layouts. In ad…LLMs have been shown to hallucinate [24] (i.e., generate false information) and may potentially identify in…uideline violations. This paper explores the potential of using LLMs to carry out heuristic evaluation au…ally. In particular, we aim to determine their performance, strengths and limitations, and how an LLM-ba… can fit into existing design practices.

To… the potential of LLMs in conducting heuristic evaluation, we built a tool that enables designers to run au… ockups and receive text-based feedback. We package this system as a plugin for Fi… too. Figure 1 illustrates the iterative usage of this plugin. The designer prototypes

We discovered several reading challenges.



Searching

4 STUDY METHOD

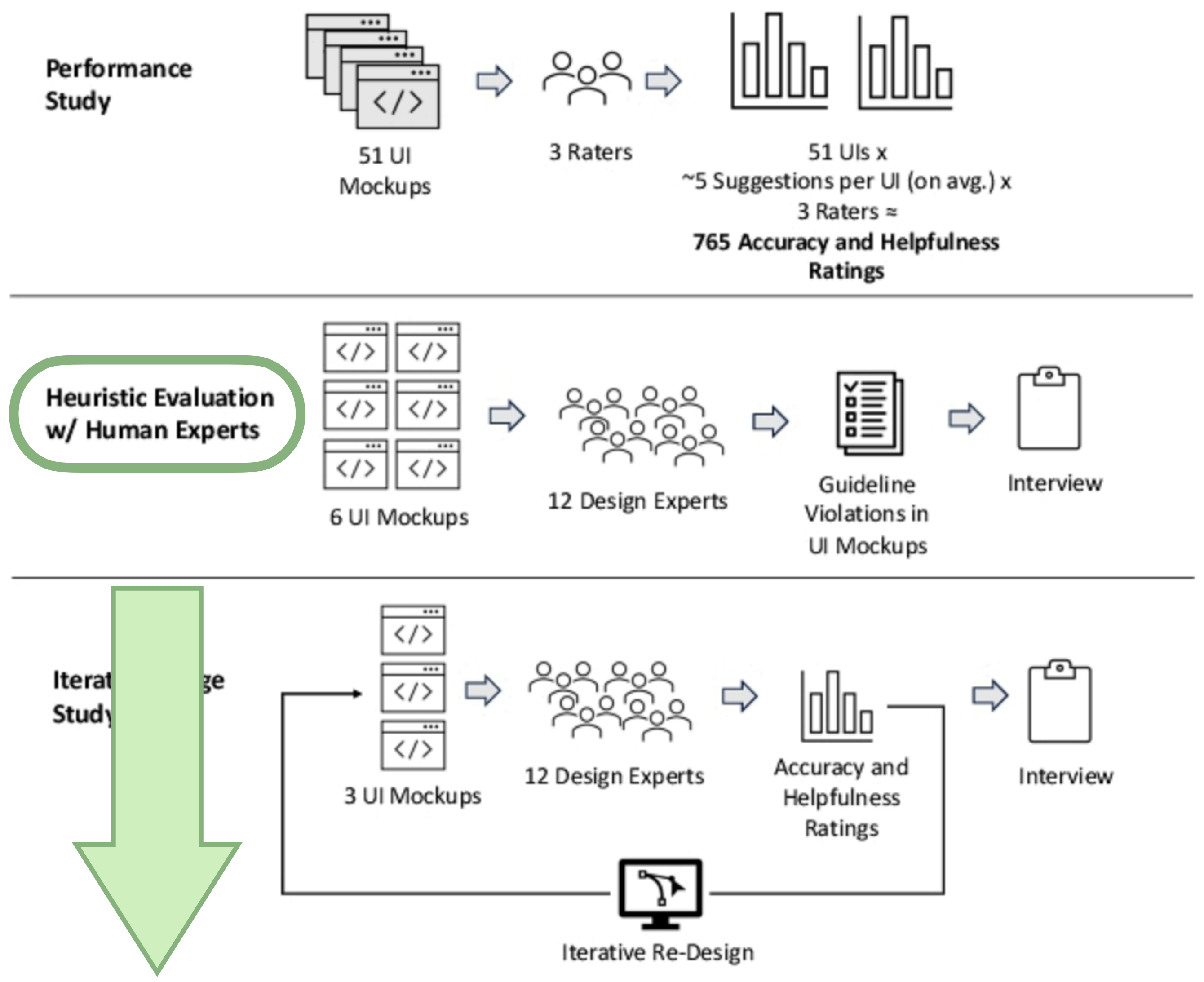
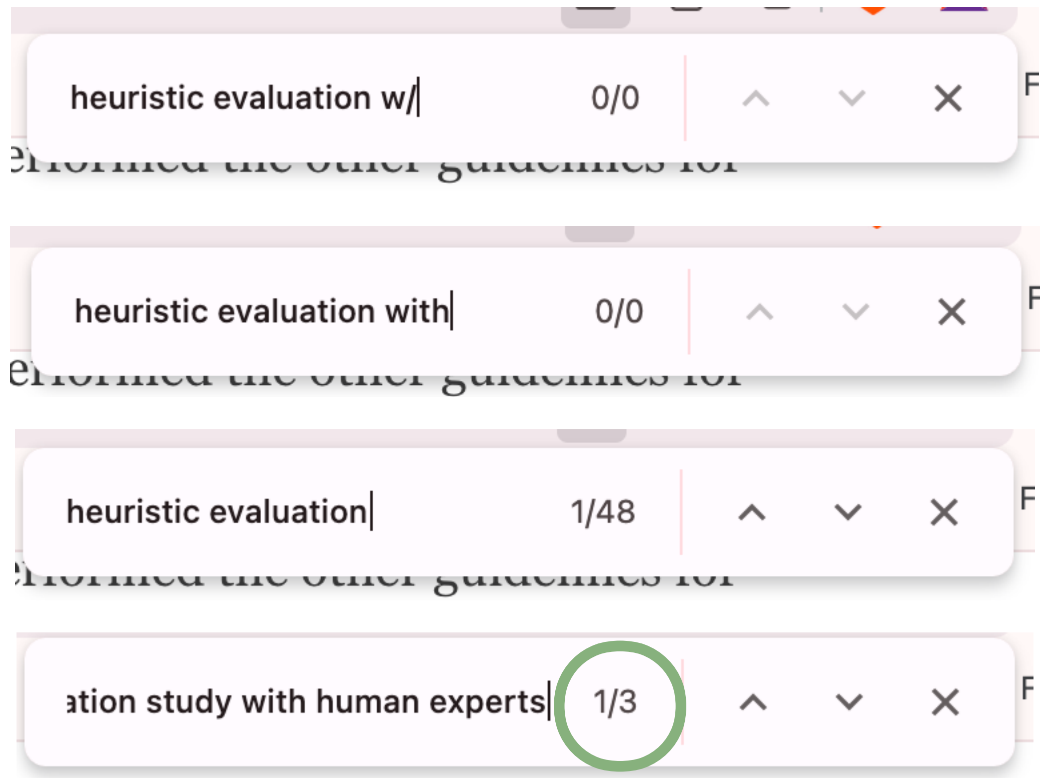
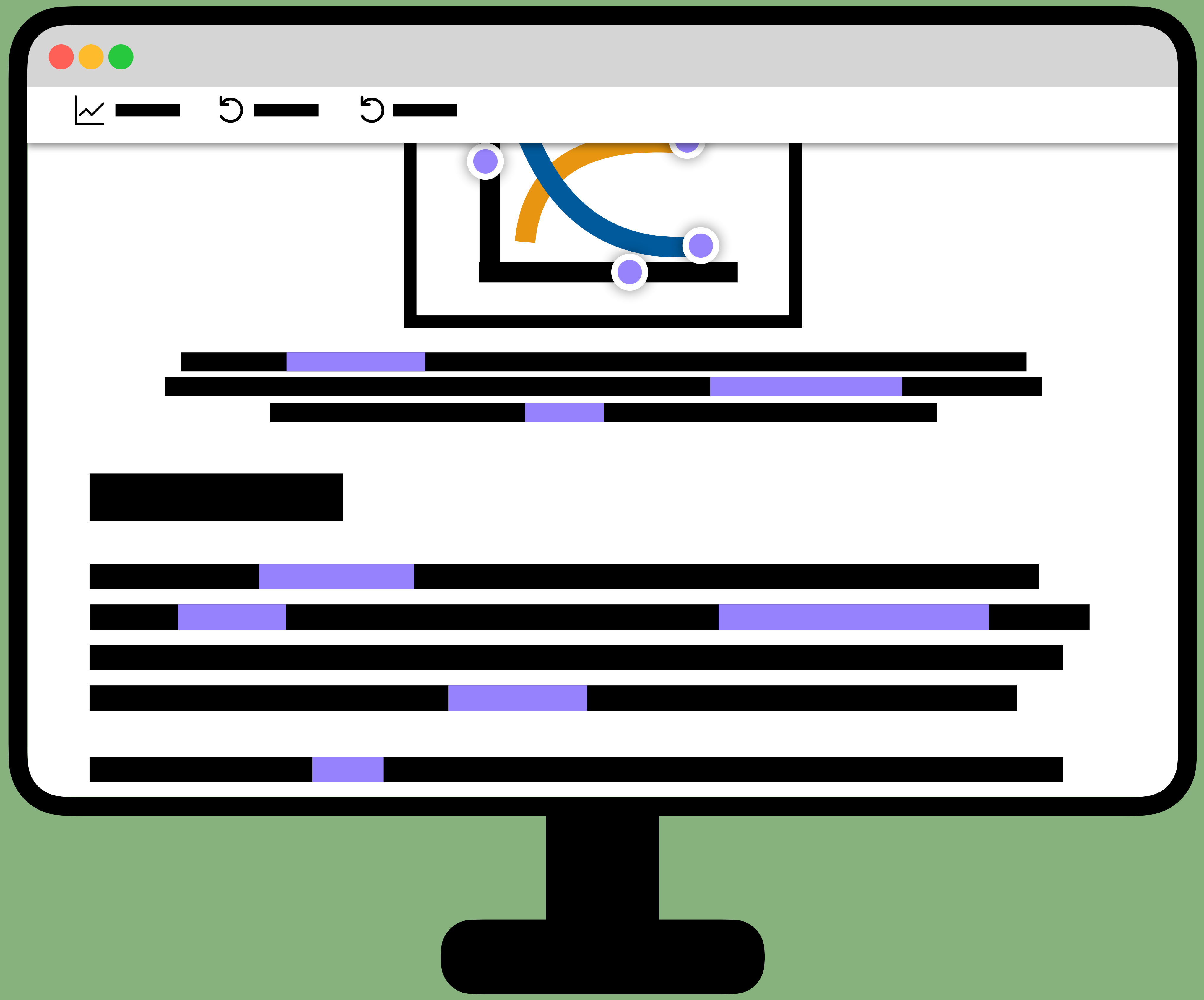


Figure 5: An illustration of the formats of the three studies. The Performance Study consists of 3 raters evaluating the accuracy and helpfulness of GPT-4-generated suggestions for 51 UI mockups. The Heuristic Evaluation Study with Human Experts consists of 12 design experts, who each looked for guideline violations in 6 UIs, and finishes with an interview asking them to compare their violations with those found by the LLM. Finally, the Iterative Usage study comprises of another group of 12 design experts, each working with 3 UI mockups. For each mockup, the expert iteratively revises the design based on the LLM's valid suggestions and rates the LLM's feedback, going through 2-3 rounds of this per UI. The Usage study concludes with an interview about the expert's experience with the tool.

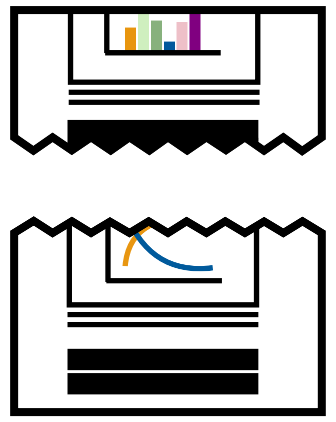


To explore the potential of GPT-4 in automating heuristic evaluation, we carried out three studies (see Figure 5).

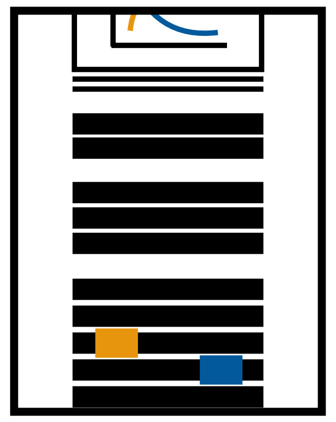
We developed a prototype of fine-grained augmentations to address these challenges.



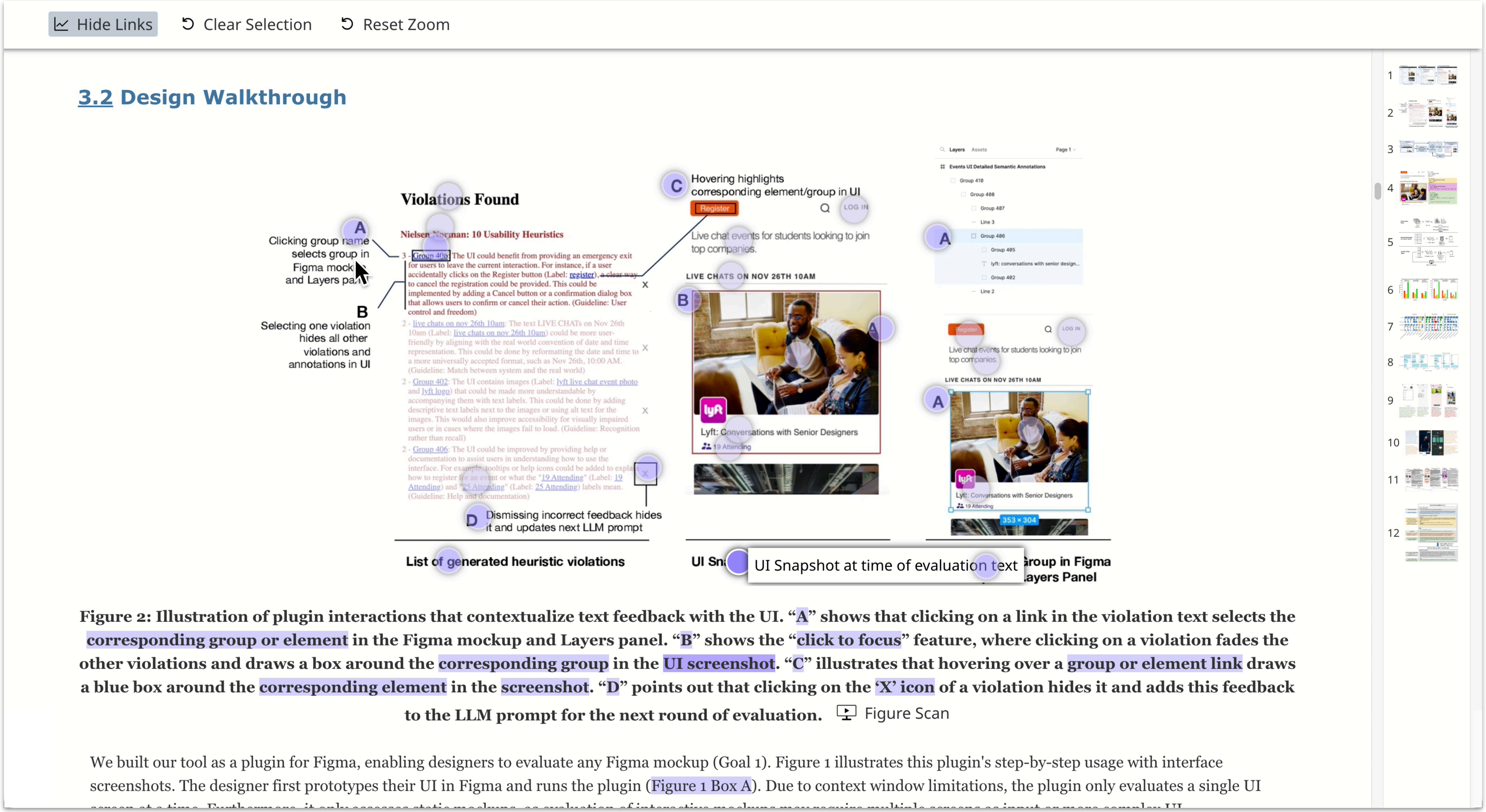
Summon more info by clicking on points or phrases.



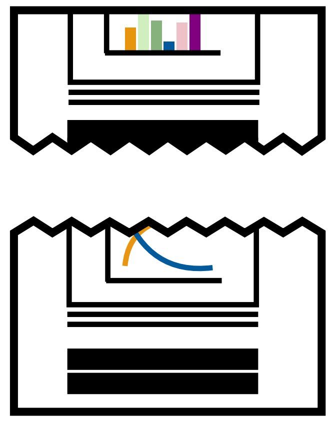
Fragmentation



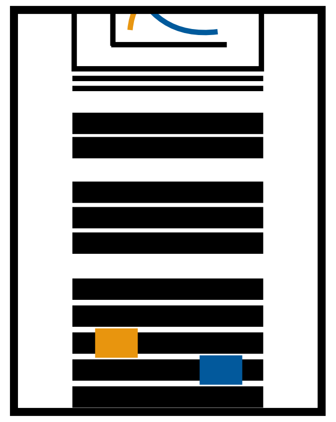
Visibility



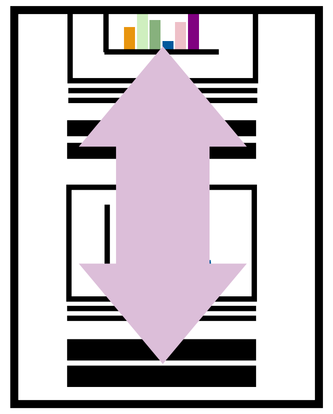
Consolidate scattered details in the reference panel.



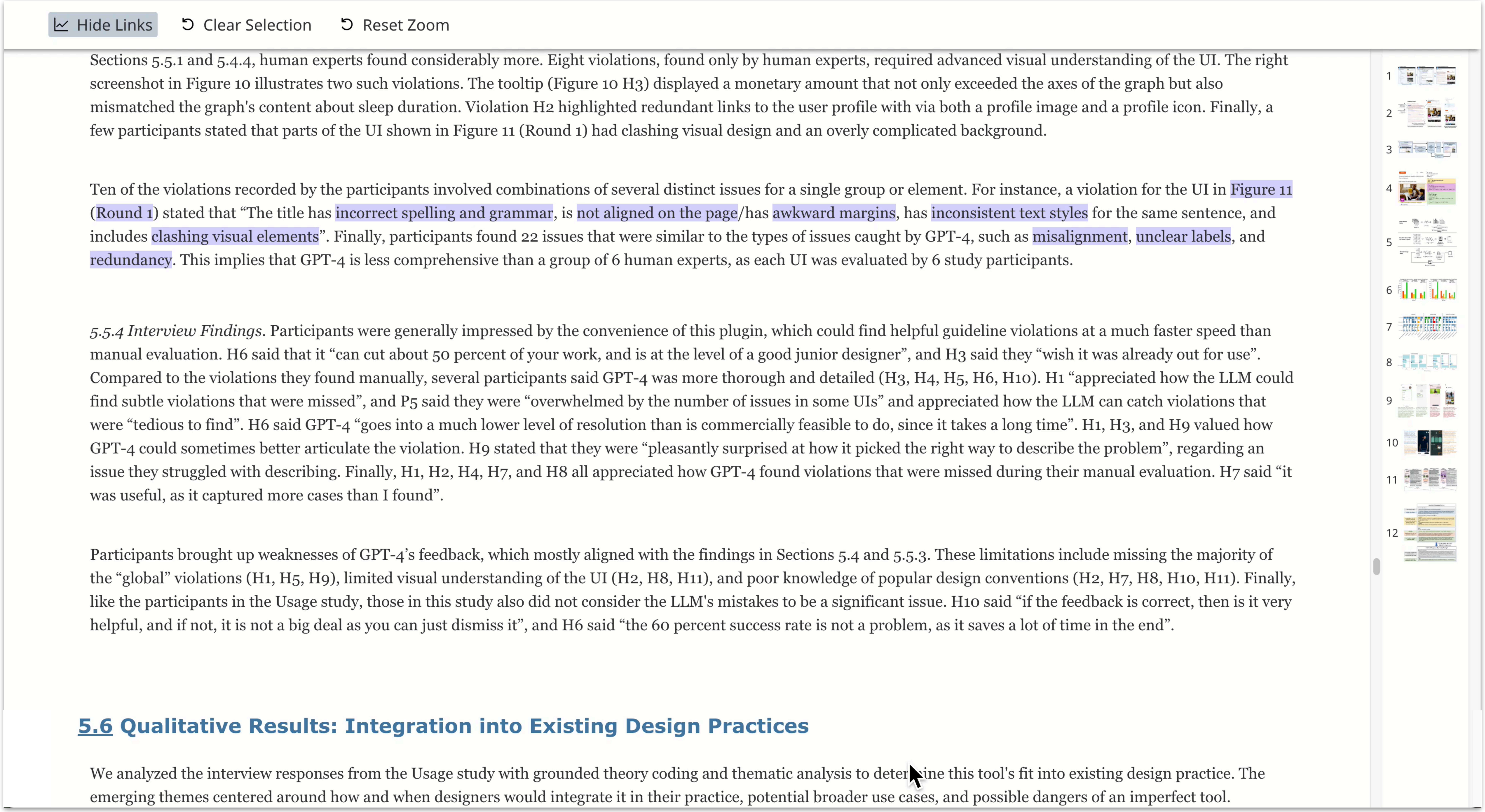
Fragmentation



Visibility



Searching



Hide Links Clear Selection Reset Zoom

Sections 5.5.1 and 5.4.4, human experts found considerably more. Eight violations, found only by human experts, required advanced visual understanding of the UI. The right screenshot in Figure 10 illustrates two such violations. The tooltip (Figure 10 H3) displayed a monetary amount that not only exceeded the axes of the graph but also mismatched the graph's content about sleep duration. Violation H2 highlighted redundant links to the user profile with via both a profile image and a profile icon. Finally, a few participants stated that parts of the UI shown in Figure 11 (Round 1) had clashing visual design and an overly complicated background.

Ten of the violations recorded by the participants involved combinations of several distinct issues for a single group or element. For instance, a violation for the UI in Figure 11 (Round 1) stated that "The title has incorrect spelling and grammar, is not aligned on the page/has awkward margins, has inconsistent text styles for the same sentence, and includes clashing visual elements". Finally, participants found 22 issues that were similar to the types of issues caught by GPT-4, such as misalignment, unclear labels, and redundancy. This implies that GPT-4 is less comprehensive than a group of 6 human experts, as each UI was evaluated by 6 study participants.

5.5.4 Interview Findings. Participants were generally impressed by the convenience of this plugin, which could find helpful guideline violations at a much faster speed than manual evaluation. H6 said that it "can cut about 50 percent of your work, and is at the level of a good junior designer", and H3 said they "wish it was already out for use". Compared to the violations they found manually, several participants said GPT-4 was more thorough and detailed (H3, H4, H5, H6, H10). H1 "appreciated how the LLM could find subtle violations that were missed", and P5 said they were "overwhelmed by the number of issues in some UIs" and appreciated how the LLM can catch violations that were "tedious to find". H6 said GPT-4 "goes into a much lower level of resolution than is commercially feasible to do, since it takes a long time". H1, H3, and H9 valued how GPT-4 could better articulate the violation. H9 stated that they were "pleasantly surprised at how it picked the right way to describe the problem", regarding an issue they struggled with describing. Finally, H1, H2, H4, H7, and H8 all appreciated how GPT-4 found violations that were missed during their manual evaluation. H7 said "it was useful, as it captured more cases than I found".

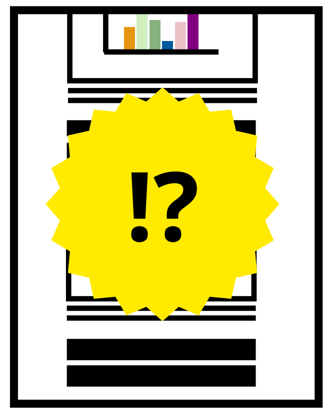
Participants brought up weaknesses of GPT-4's feedback, which mostly aligned with the findings in Sections 5.4 and 5.5.3. These limitations include missing the majority of the "global" violations (H1, H5, H9), limited visual understanding of the UI (H2, H8, H11), and poor knowledge of popular design conventions (H2, H7, H8, H10, H11). Finally, like the participants in the Usage study, those in this study also did not consider the LLM's mistakes to be a significant issue. H10 said "if the feedback is correct, then is it very helpful, and if not, it is not a big deal as you can just dismiss it", and H6 said "the 60 percent success rate is not a problem, as it saves a lot of time in the end".

5.6 Qualitative Results: Integration into Existing Design Practices

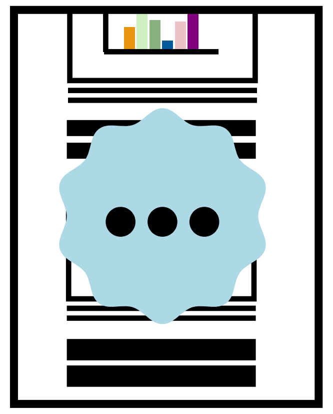
We analyzed the interview responses from the Usage study with grounded theory coding and thematic analysis to determine this tool's fit into existing design practice. The emerging themes centered around how and when designers would integrate it in their practice, potential broader use cases, and possible dangers of an imperfect tool.

1
2
3
4
5
6
7
8
9
10
11
12

Read figures one step at a time with figure scans.



Complexity



Interpretation

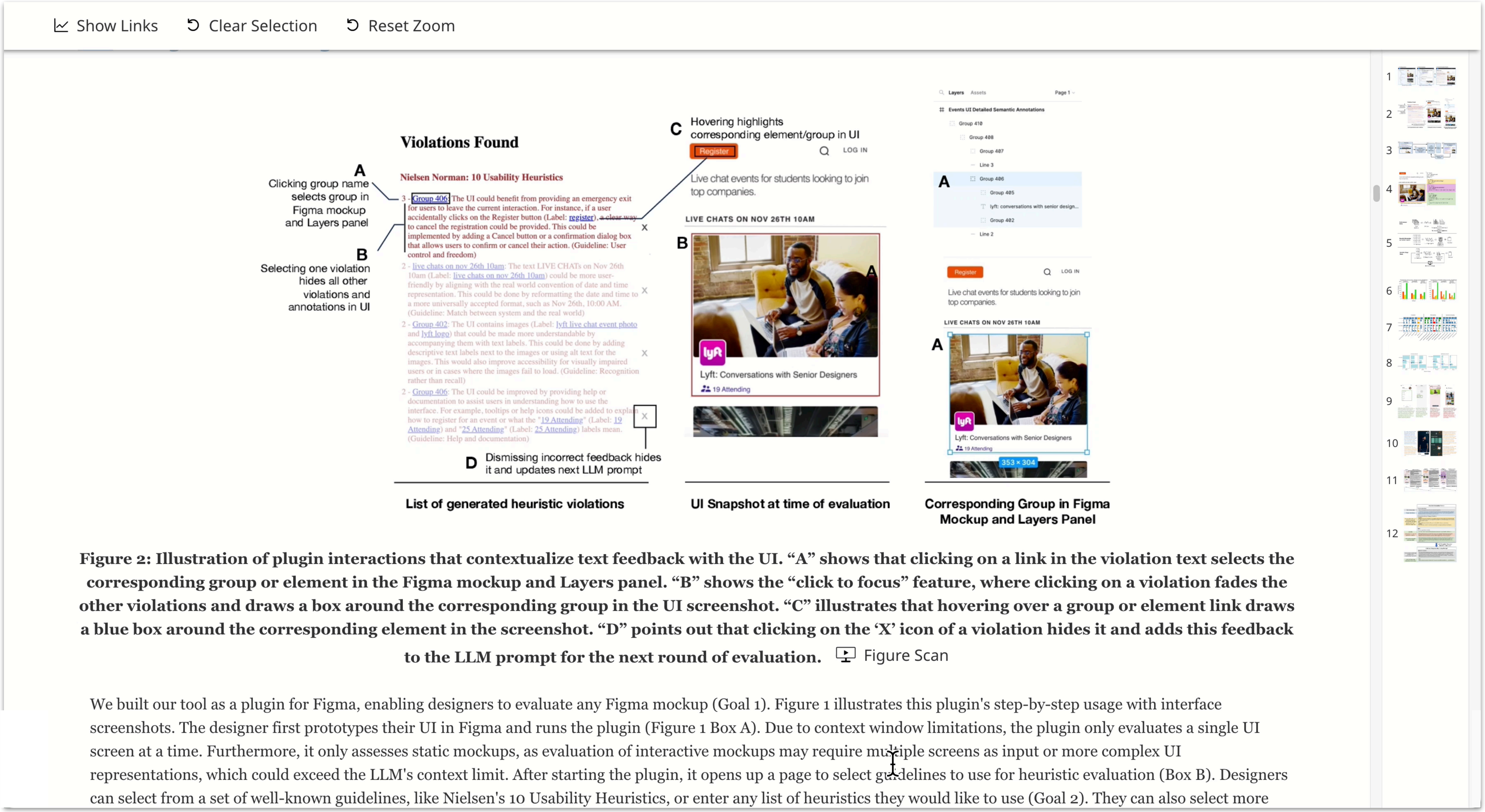
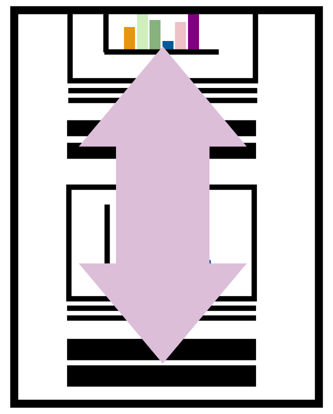

Show Links Clear Selection Reset Zoom

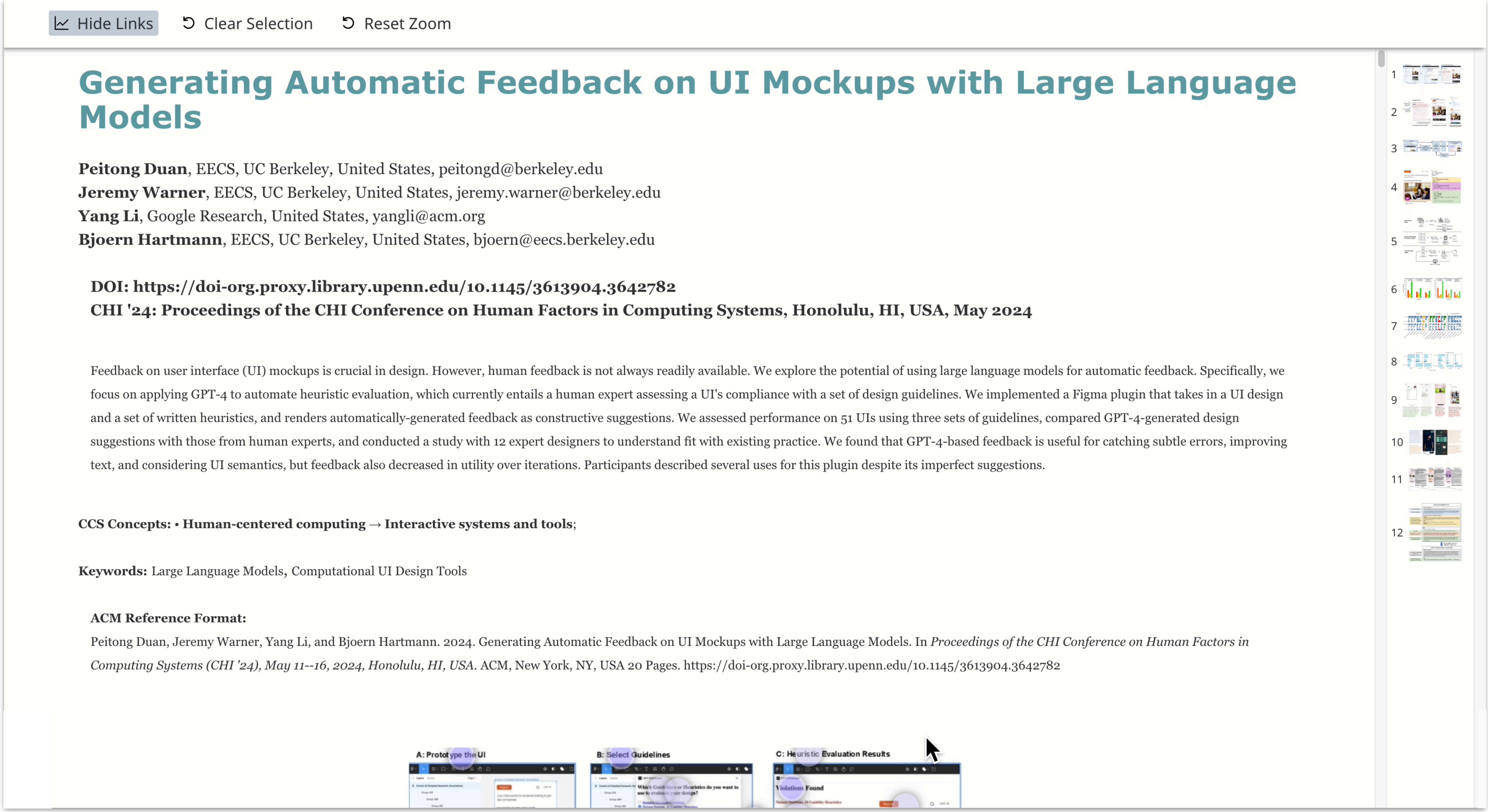
Figure 2: Illustration of plugin interactions that contextualize text feedback with the UI. "A" shows that clicking on a link in the violation text selects the corresponding group or element in the Figma mockup and Layers panel. "B" shows the "click to focus" feature, where clicking on a violation fades the other violations and draws a box around the corresponding group in the UI screenshot. "C" illustrates that hovering over a group or element link draws a blue box around the corresponding element in the screenshot. "D" points out that clicking on the 'X' icon of a violation hides it and adds this feedback to the LLM prompt for the next round of evaluation. Figure Scan

We built our tool as a plugin for Figma, enabling designers to evaluate any Figma mockup (Goal 1). Figure 1 illustrates this plugin's step-by-step usage with interface screenshots. The designer first prototypes their UI in Figma and runs the plugin (Figure 1 Box A). Due to context window limitations, the plugin only evaluates a single UI screen at a time. Furthermore, it only assesses static mockups, as evaluation of interactive mockups may require multiple screens as input or more complex UI representations, which could exceed the LLM's context limit. After starting the plugin, it opens up a page to select guidelines to use for heuristic evaluation (Box B). Designers can select from a set of well-known guidelines, like Nielsen's 10 Usability Heuristics, or enter any list of heuristics they would like to use (Goal 2). They can also select more

Jump directly to specific figures with the visual index.



Searching



Generating Automatic Feedback on UI Mockups with Large Language Models

Peitong Duan, EECS, UC Berkeley, United States, peitongd@berkeley.edu
Jeremy Warner, EECS, UC Berkeley, United States, jeremy.warner@berkeley.edu
Yang Li, Google Research, United States, yangli@acm.org
Bjoern Hartmann, EECS, UC Berkeley, United States, bjoern@eecs.berkeley.edu

DOI: https://doi-org.proxy.library.upenn.edu/10.1145/3613904.3642782
CHI '24: Proceedings of the CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, May 2024

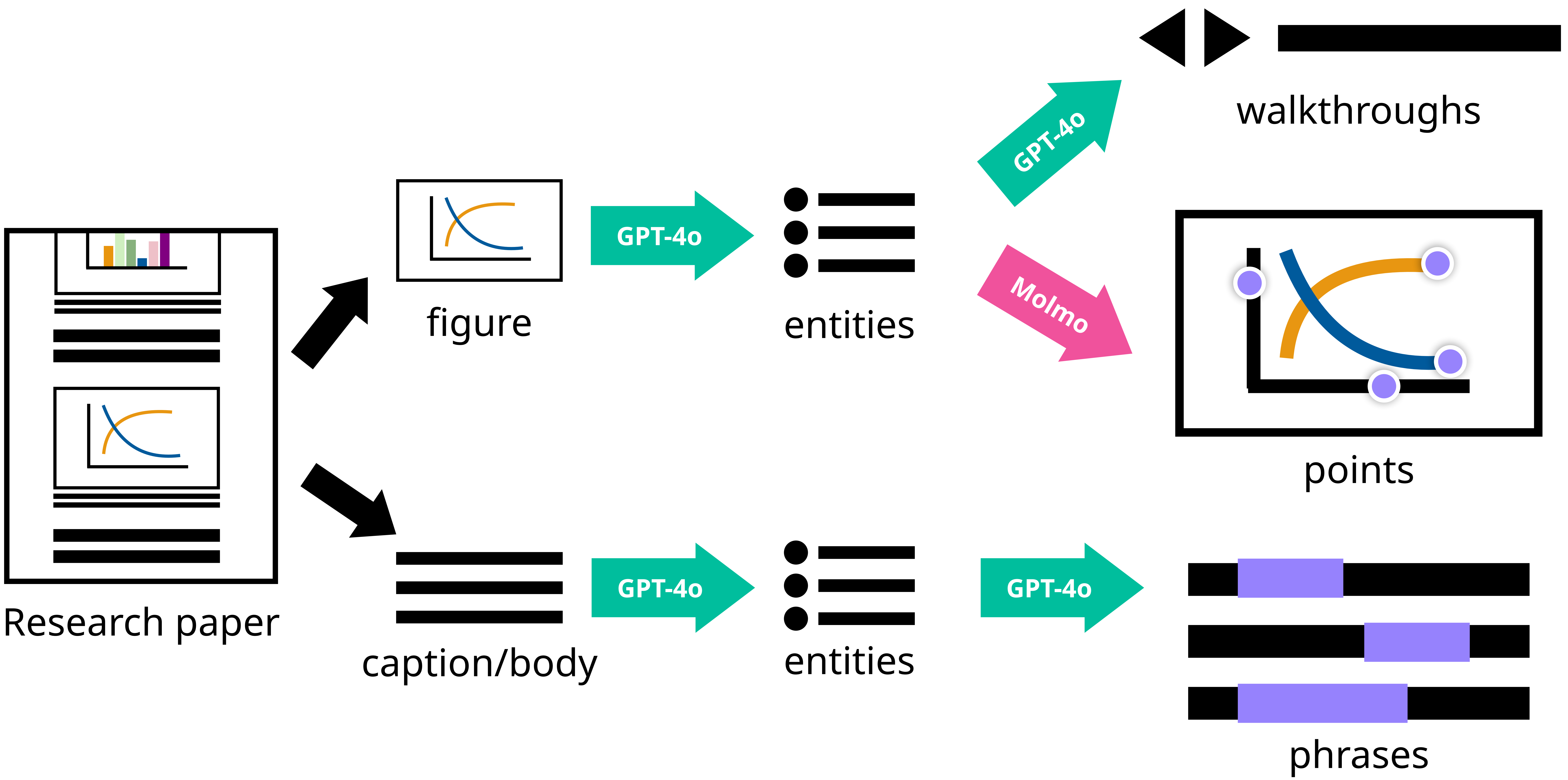
Feedback on user interface (UI) mockups is crucial in design. However, human feedback is not always readily available. We explore the potential of using large language models for automatic feedback. Specifically, we focus on applying GPT-4 to automate heuristic evaluation, which currently entails a human expert assessing a UI's compliance with a set of design guidelines. We implemented a Figma plugin that takes in a UI design and a set of written heuristics, and renders automatically-generated feedback as constructive suggestions. We assessed performance on 51 UIs using three sets of guidelines, compared GPT-4-generated design suggestions with those from human experts, and conducted a study with 12 expert designers to understand fit with existing practice. We found that GPT-4-based feedback is useful for catching subtle errors, improving text, and considering UI semantics, but feedback also decreased in utility over iterations. Participants described several uses for this plugin despite its imperfect suggestions.

CCS Concepts: • **Human-centered computing** → **Interactive systems and tools**;

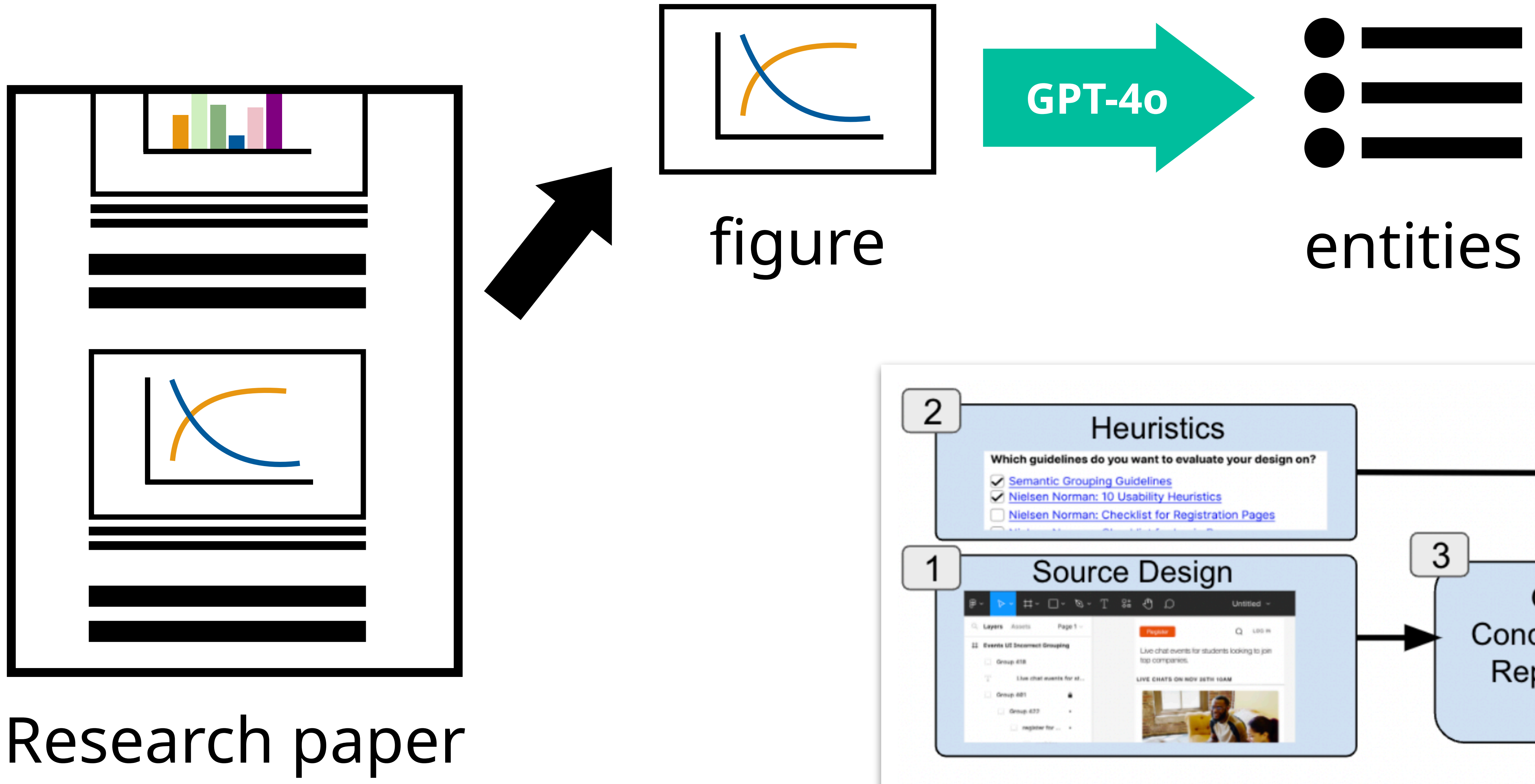
Keywords: Large Language Models, Computational UI Design Tools

ACM Reference Format:
Peitong Duan, Jeremy Warner, Yang Li, and Bjoern Hartmann. 2024. Generating Automatic Feedback on UI Mockups with Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11--16, 2024, Honolulu, HI, USA*. ACM, New York, NY, USA 20 Pages. https://doi-org.proxy.library.upenn.edu/10.1145/3613904.3642782

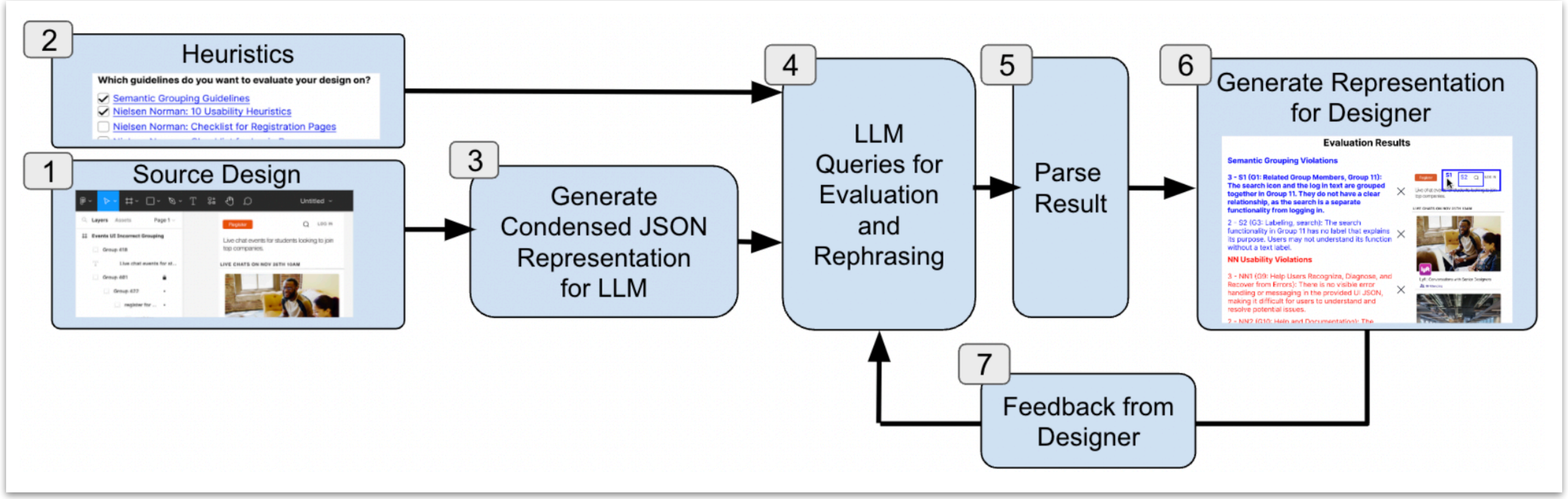
The features in our prototype were generated with a backend AI pipeline.



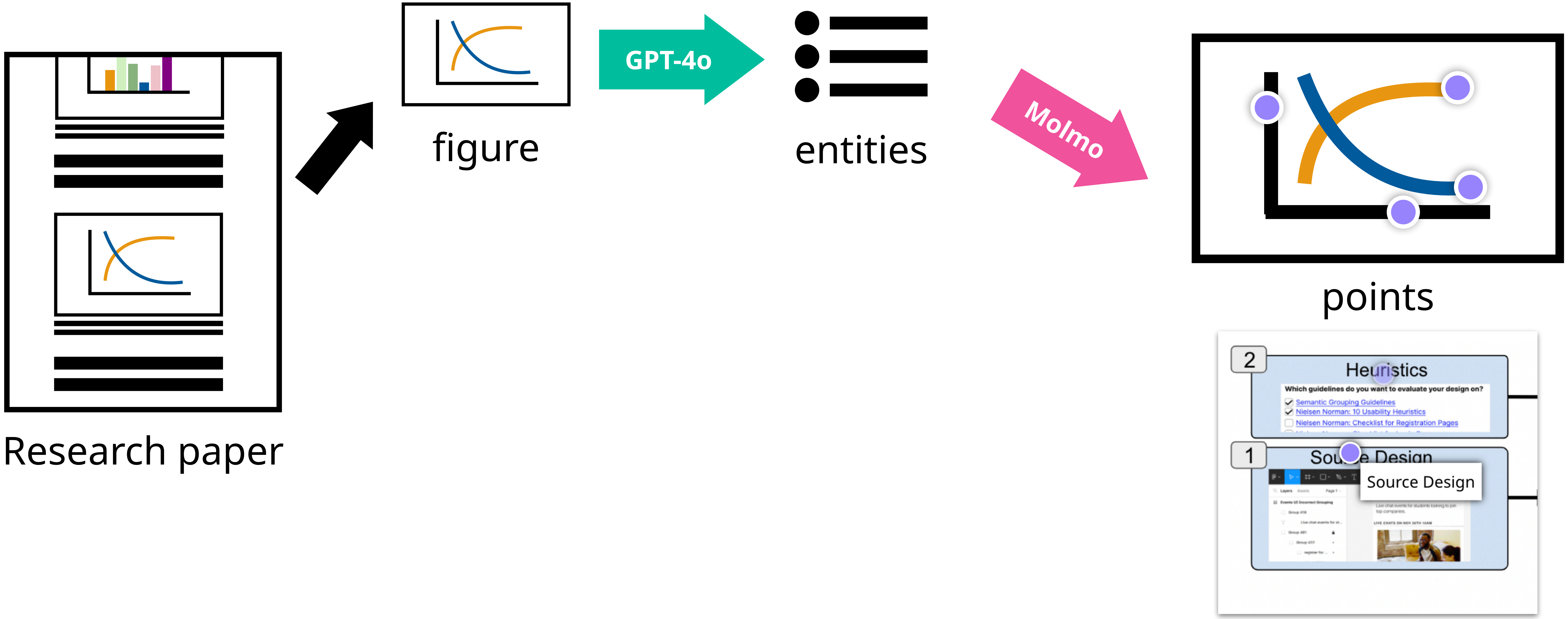
We identified figure entities with GPT-4o.



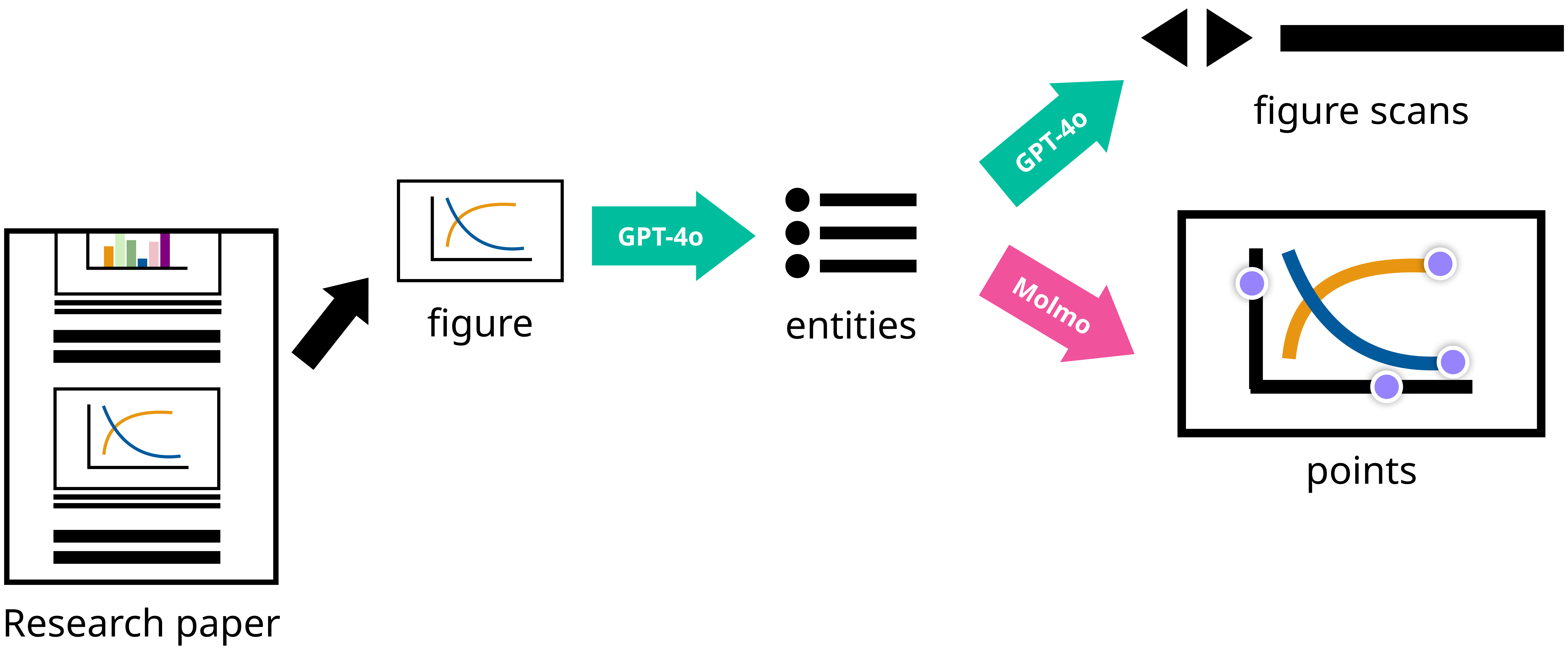
- Source Design
- Heuristics
- Parse Result
- …



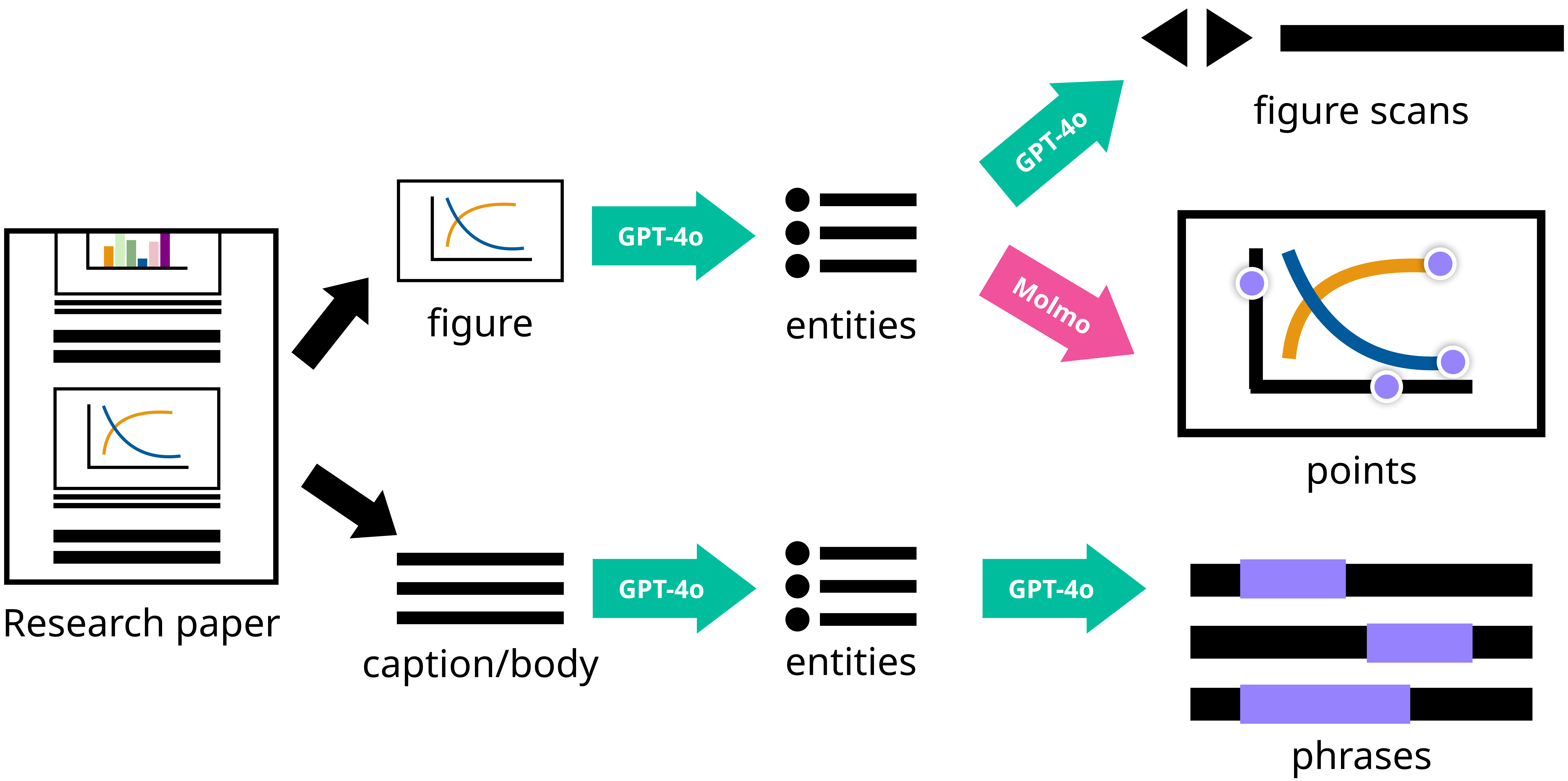
We then had Molmo mark the entities on the figures.



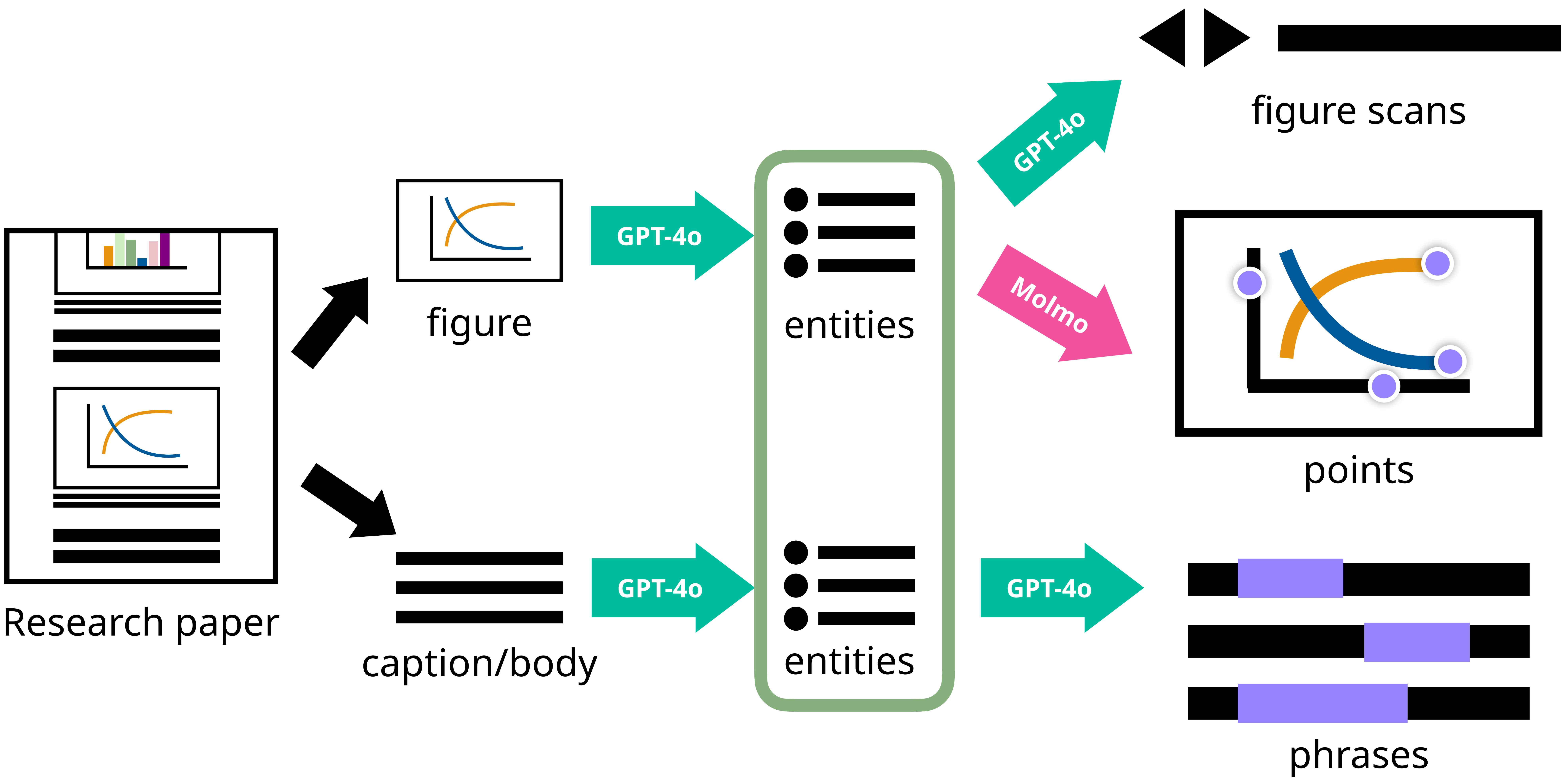
We generated descriptions of each entity with GPT-4o.



We identified important phrases with GPT-4o.



We connected points and phrases through **matching entities**.



We connected points and phrases through **matching entities**.

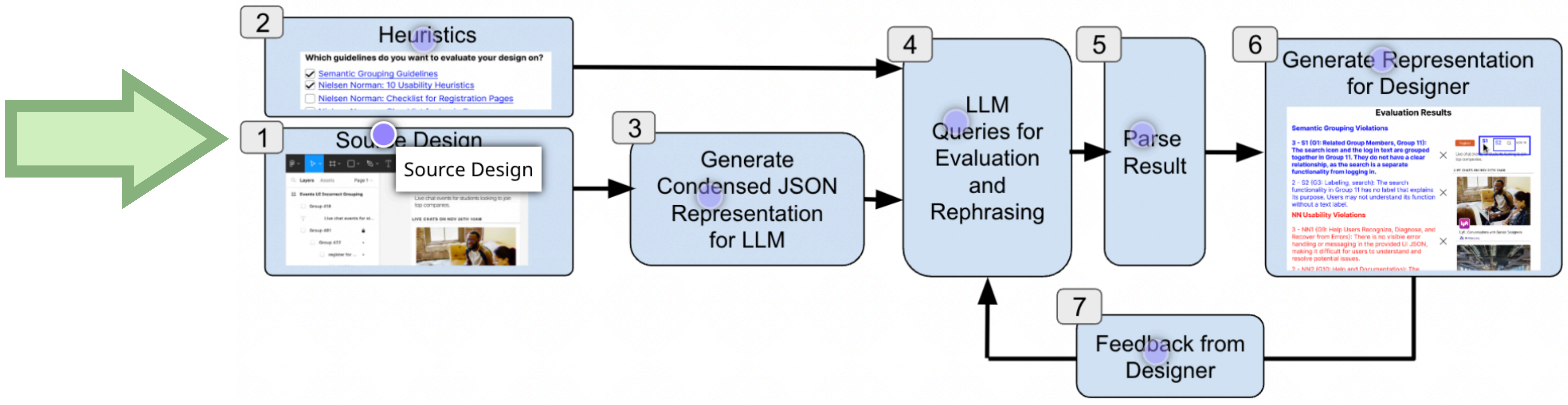
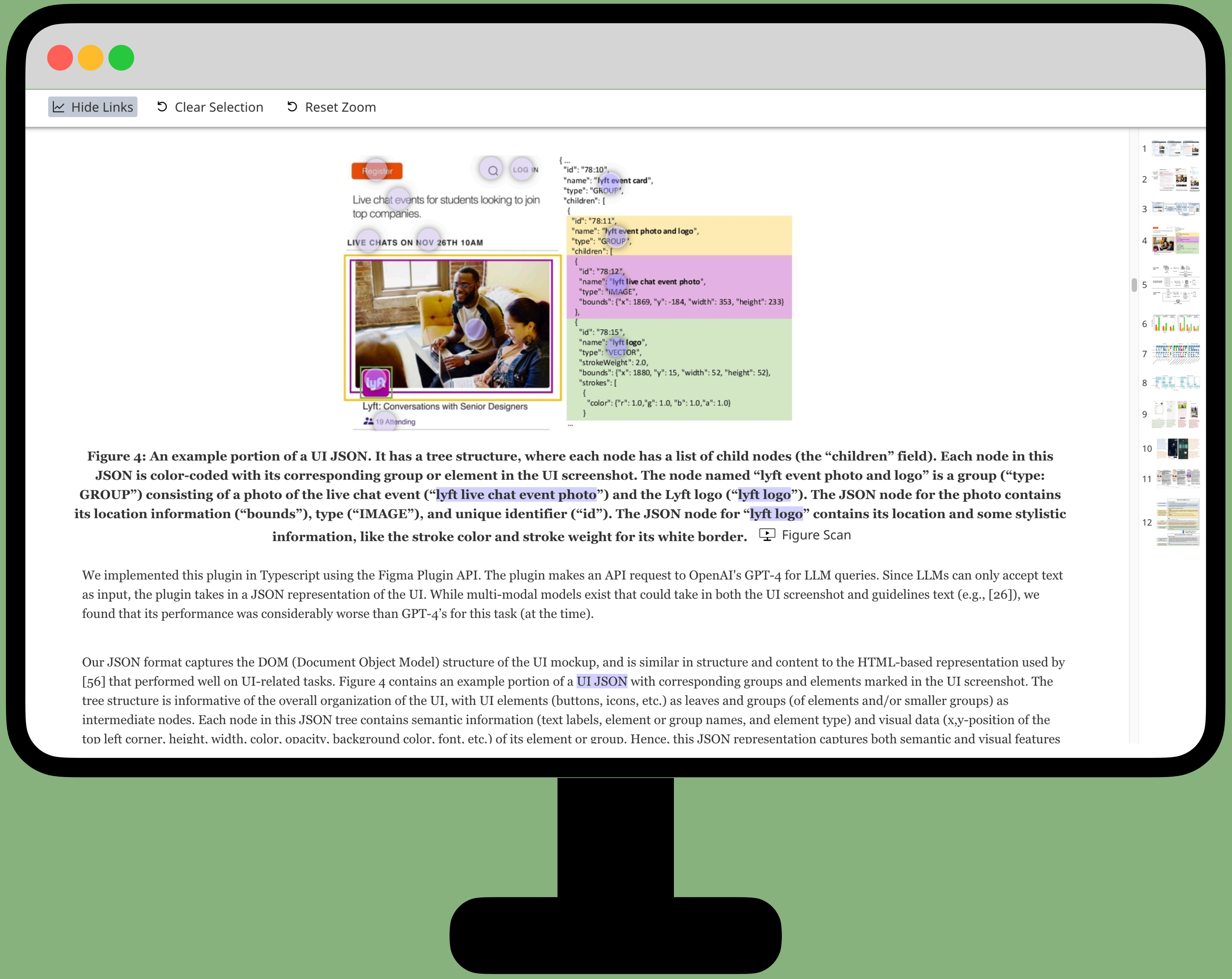


Figure 3: Our LLM-based plugin system architecture. The designer prototypes a UI in Figma (Box 1), and generates a UI representation to send to an LLM (3). The designer also selects heuristics/guidelines to use for evaluating the prototype (2), and a prompt containing the UI representation (in JSON) and guidelines is created and sent to the LLM (4). After identifying all the guideline violations, another LLM query is made to rephrase the guideline violations into constructive design advice (4). The LLM response is then programmatically parsed (5), and the plugin produces an interpretable representation of the response to display (6). The designer dismisses incorrect suggestions, which are incorporated in the LLM prompt for the next round of evaluation, if there is room in the context window (7).

phrases

Case study of prototype outcome



Fine-grained augmented reading interface based on formative study

AI pipeline for generating entities, points, phrases, and descriptions

In preparation

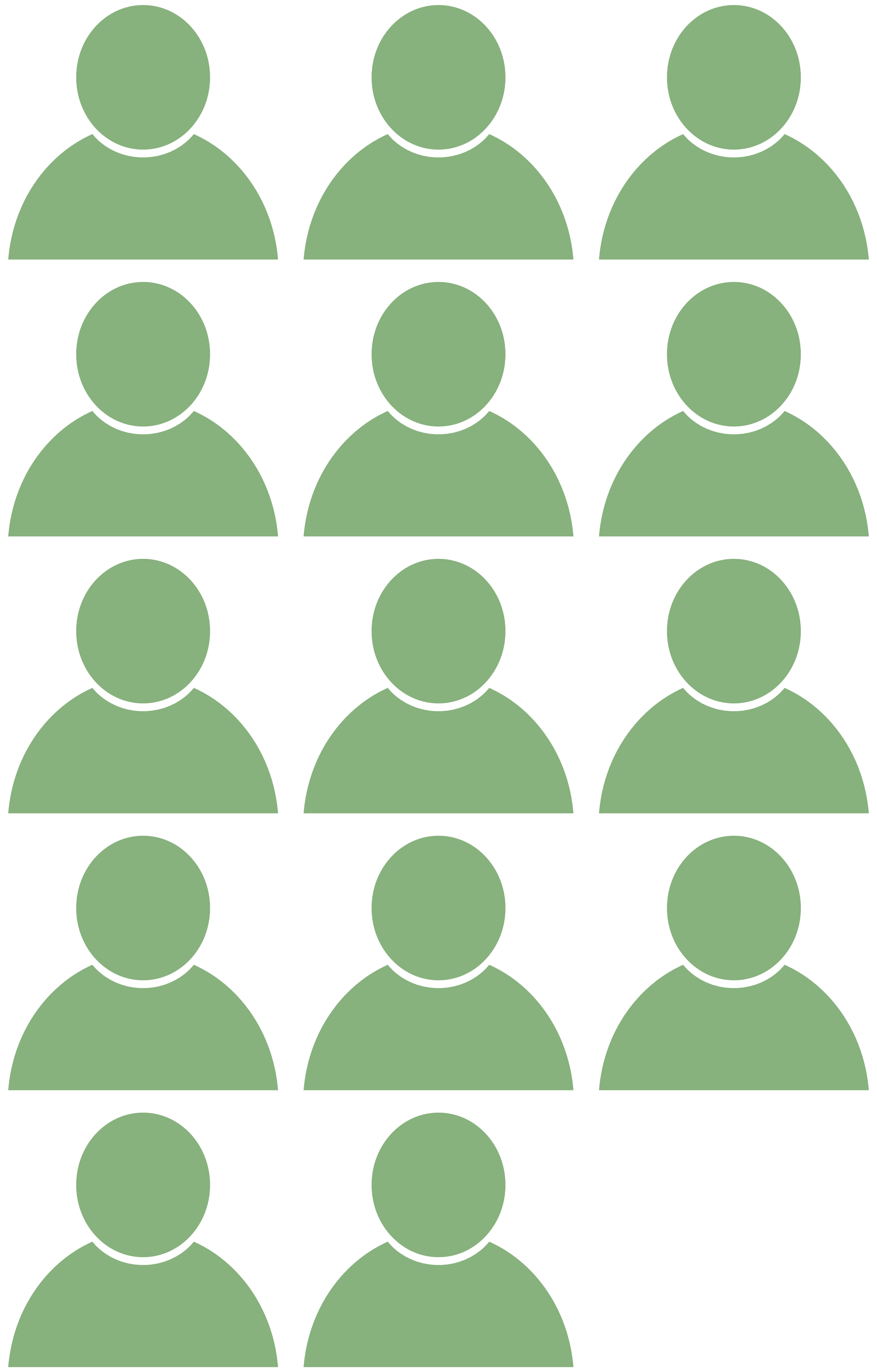
3. Behavioral analysis

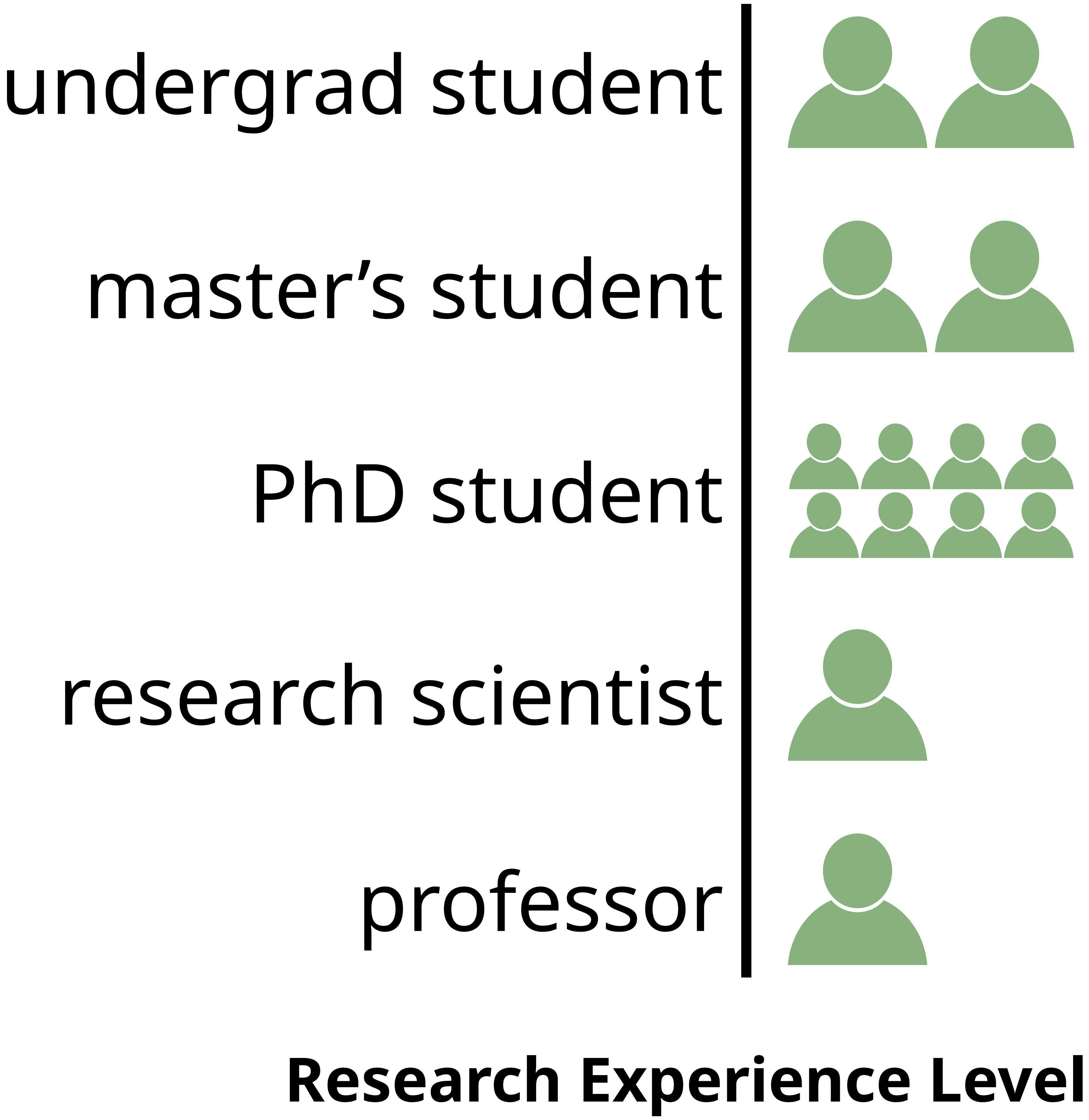
Motivation: understand effect of fine-grained augmentations

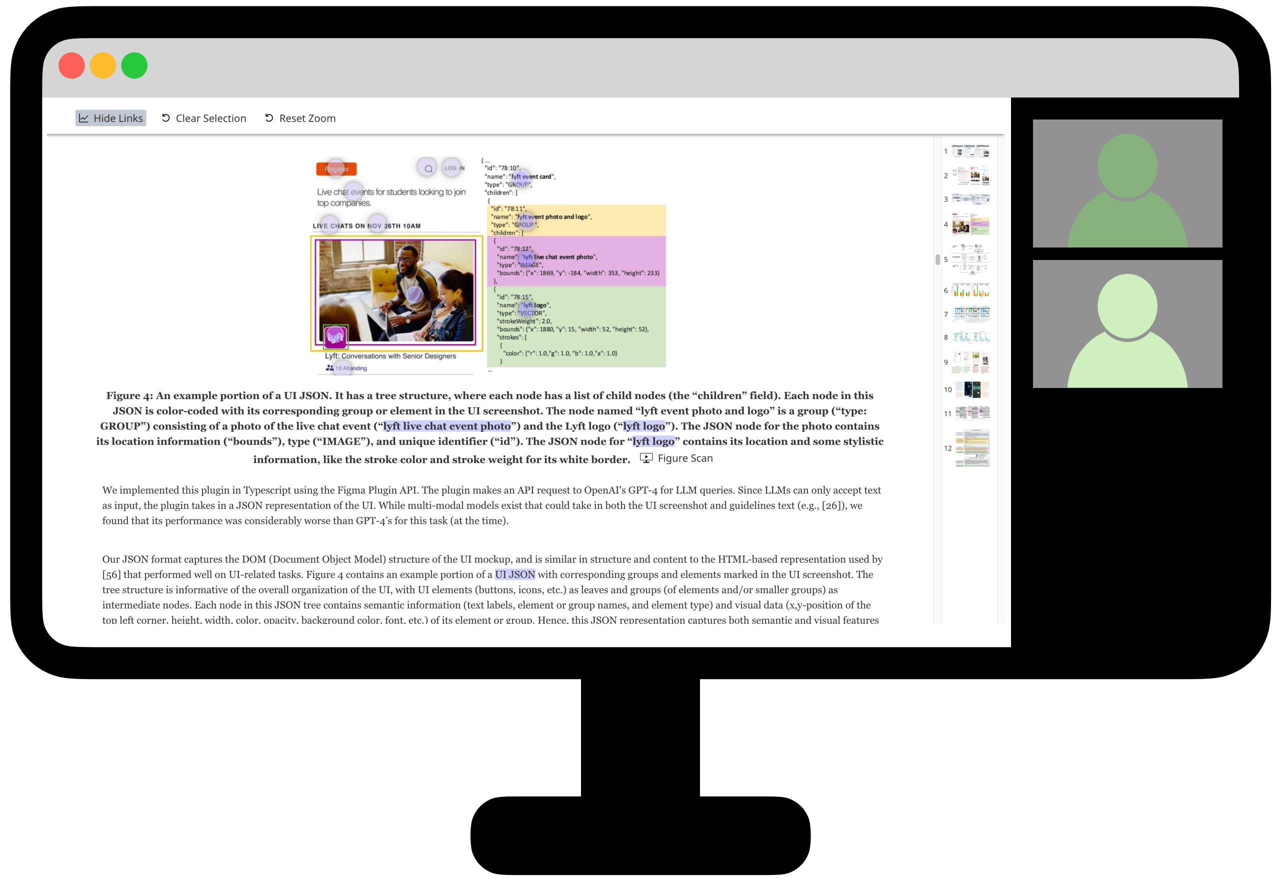
Method: observe reading with previously developed prototype

Outcome: analysis of strengths and weaknesses of augmentations

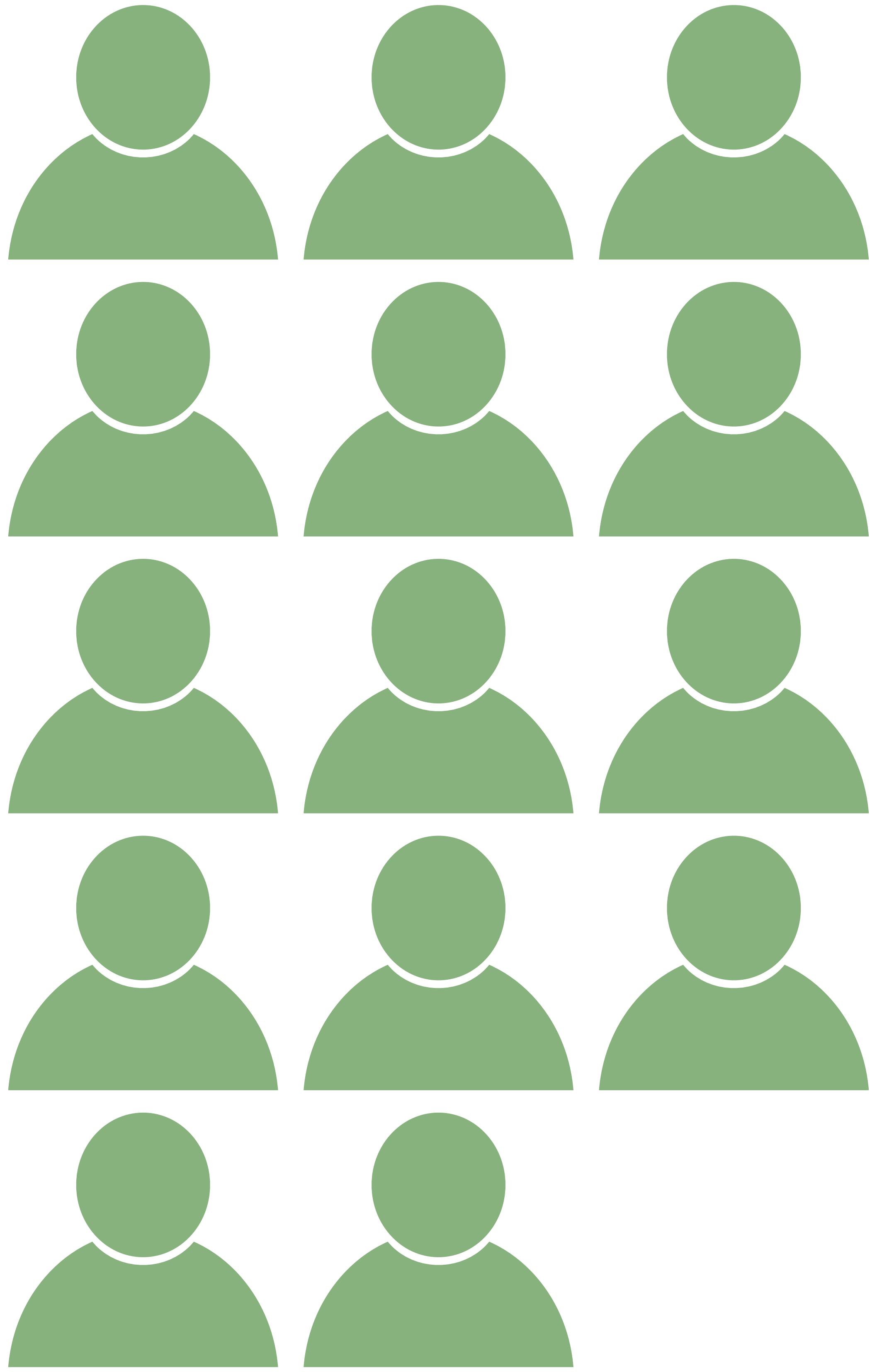
We recruited 14 HCI researchers to read a research paper with our prototype on Zoom.

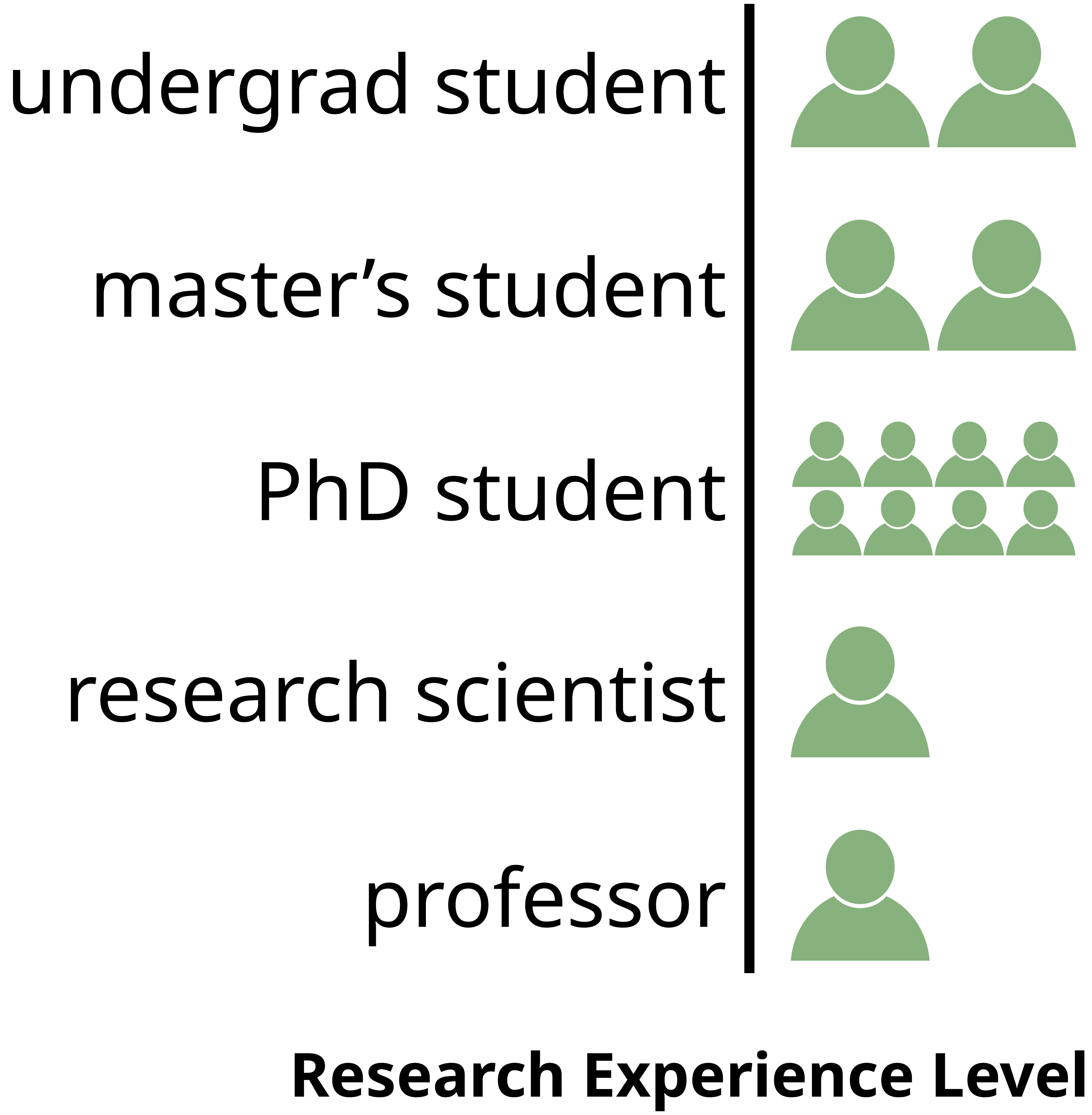






We recruited 14 HCI researchers to read a research paper with our prototype on Zoom.





Single paper due to technical limitations

Avoid interference from lack of domain knowledge

Smaller number of participants to prioritize depth of observations

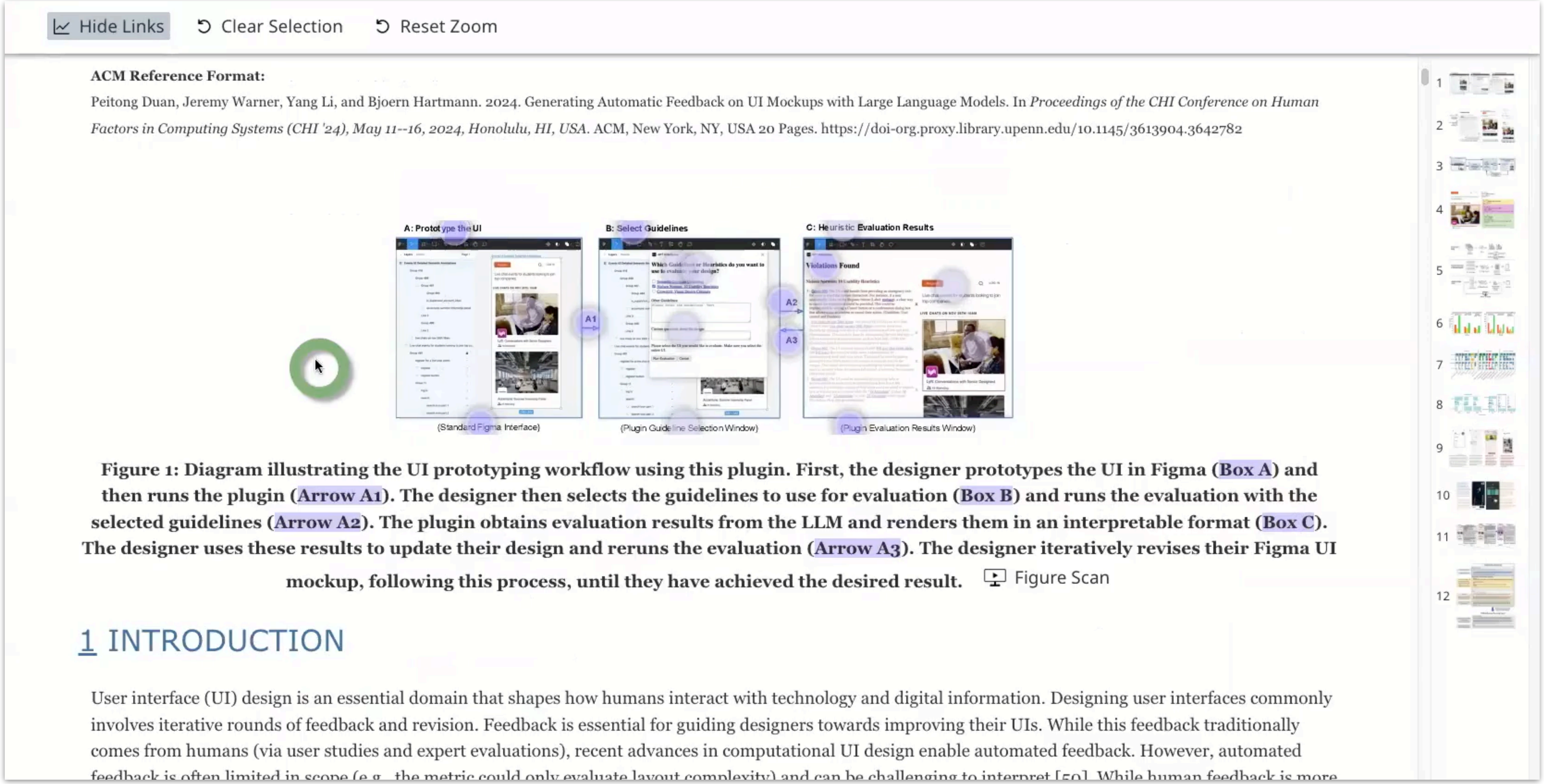
1-hour session: think-aloud reading activity + semi-structured interview

Participants used augmentations in several ways.

- Open exploration through bidirectional links
- Structured guidance through figure scans
- Reduction of search effort through reference panel

Participants used augmentations in several ways.

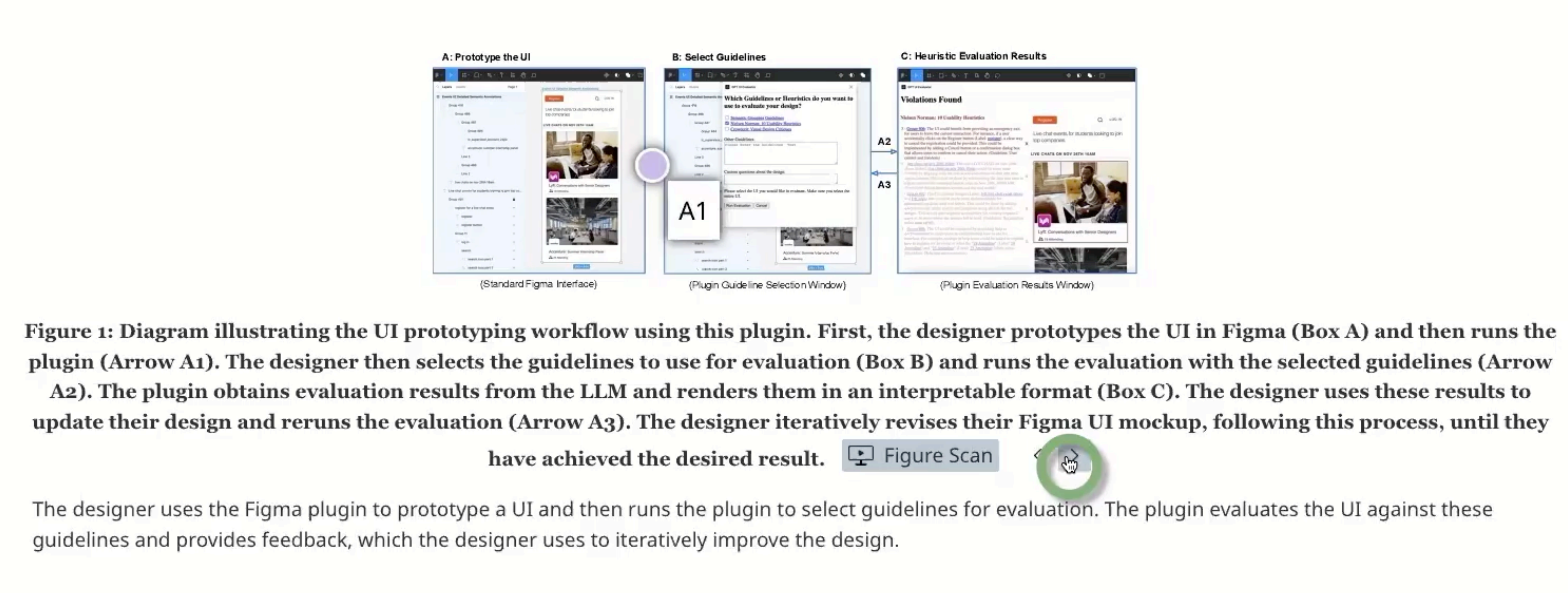
- Open exploration through bidirectional links



Participant 6, sped up 400%

Participants used augmentations in several ways.

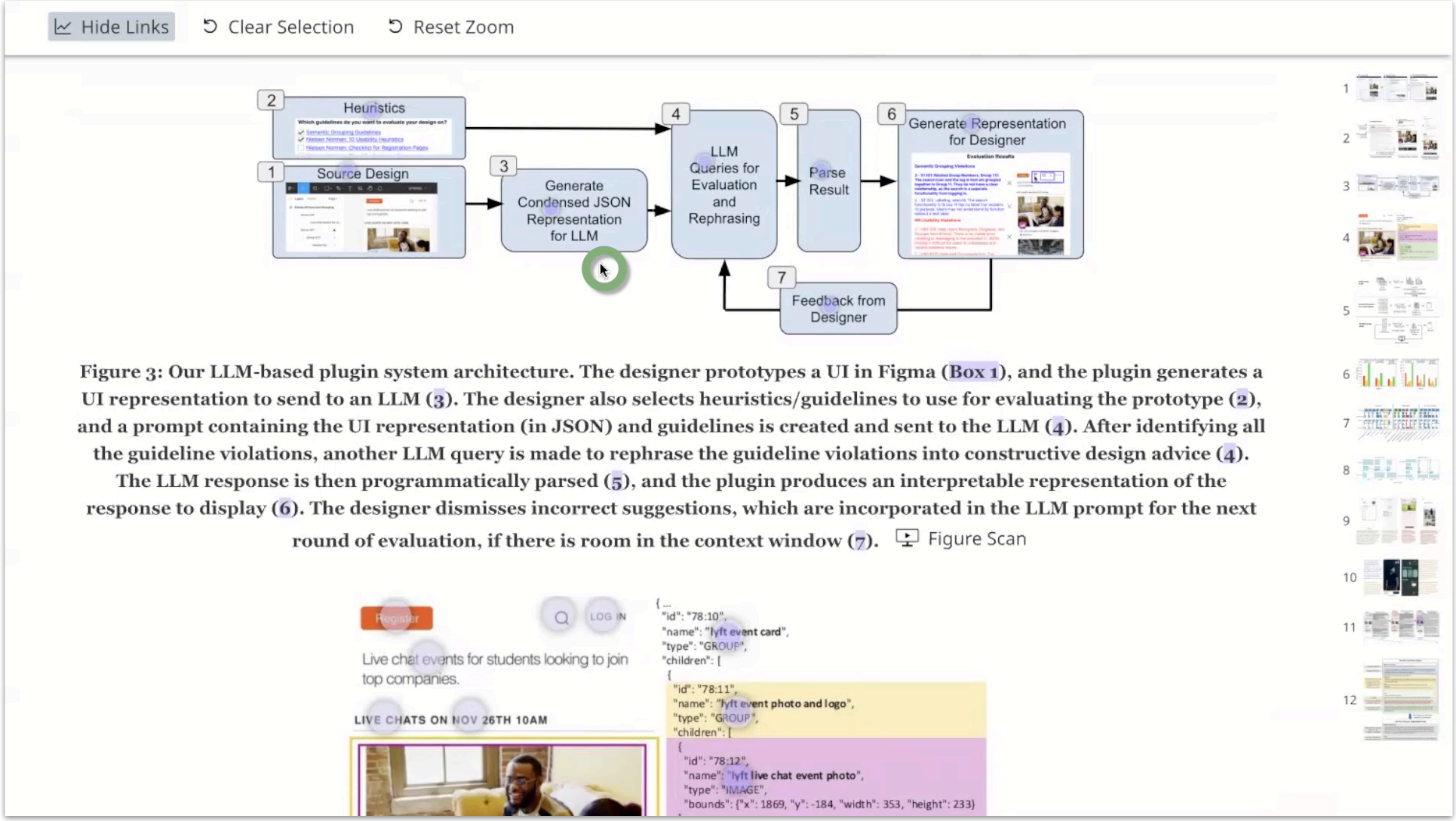
- Structured guidance through figure scans



Participant 11, sped up 300%

Participants used augmentations in several ways.

- Reduction of search effort through reference panel

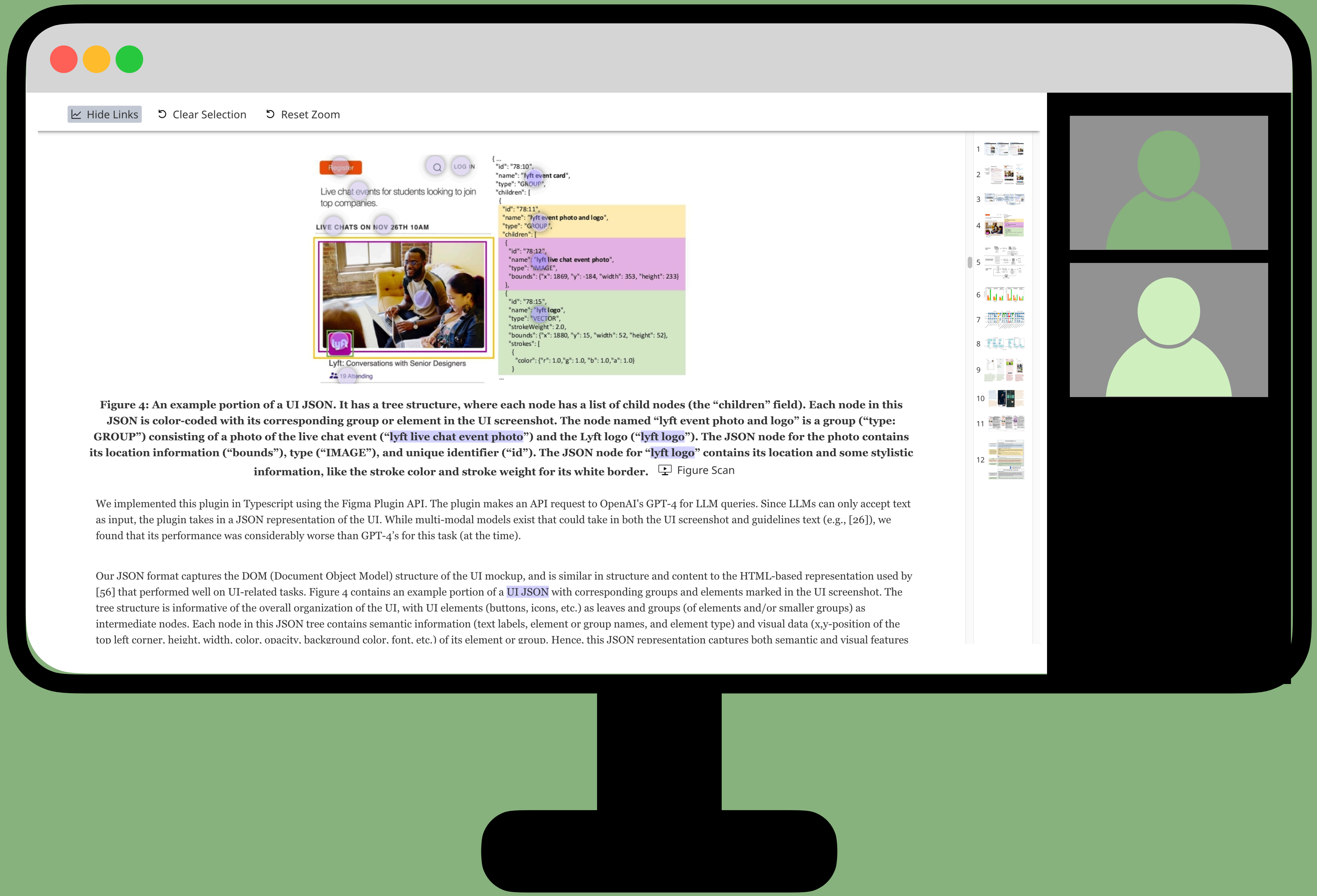


The use of augmentations presented figure-centric reading patterns.

"honeybee path": exploring multiple points at a time for the same figure

"slow dive": inspecting many parts of a figure before moving on

3. Behavioral analysis outcome



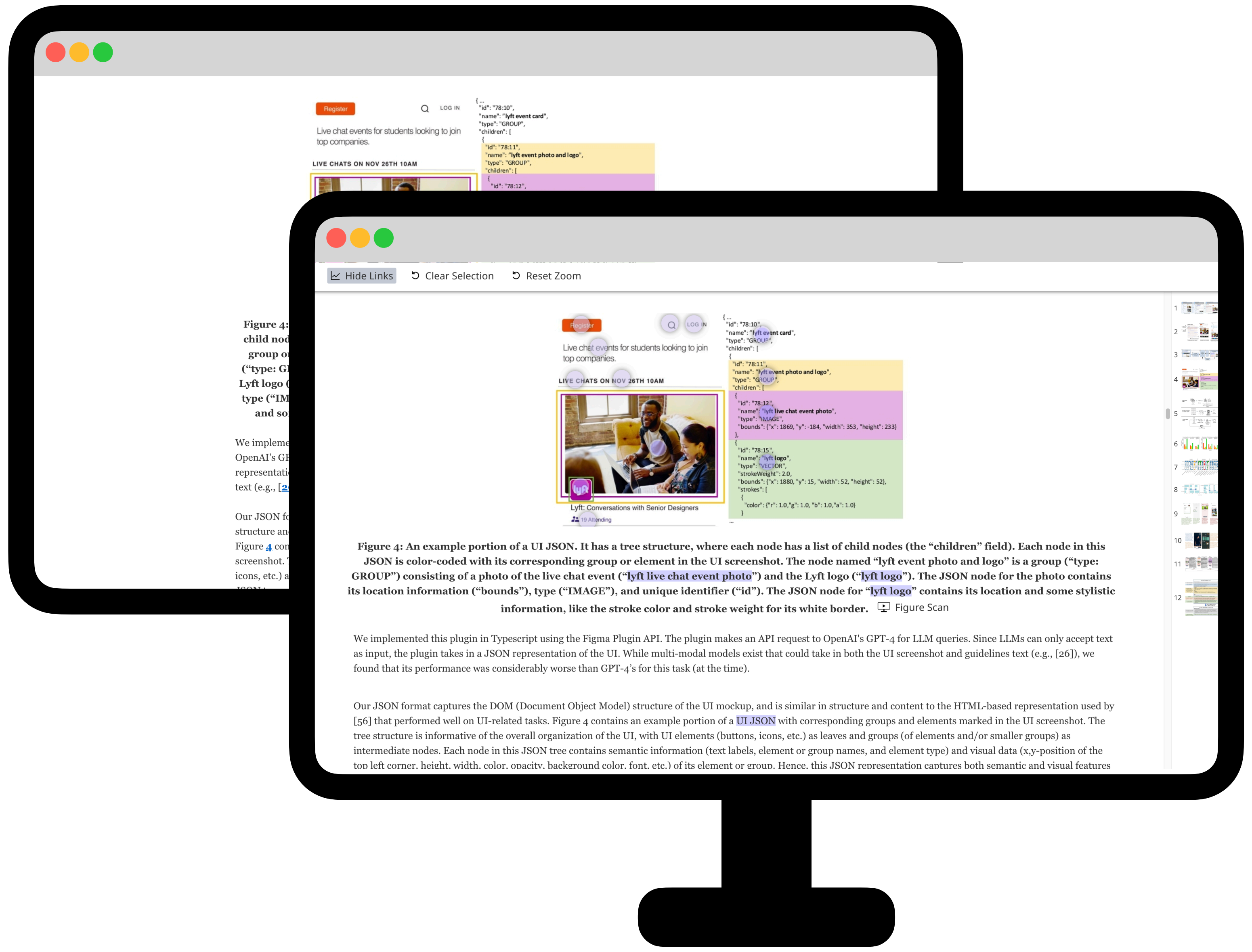
Fine-grained augmentations supported open exploration, structured guidance, faster searching

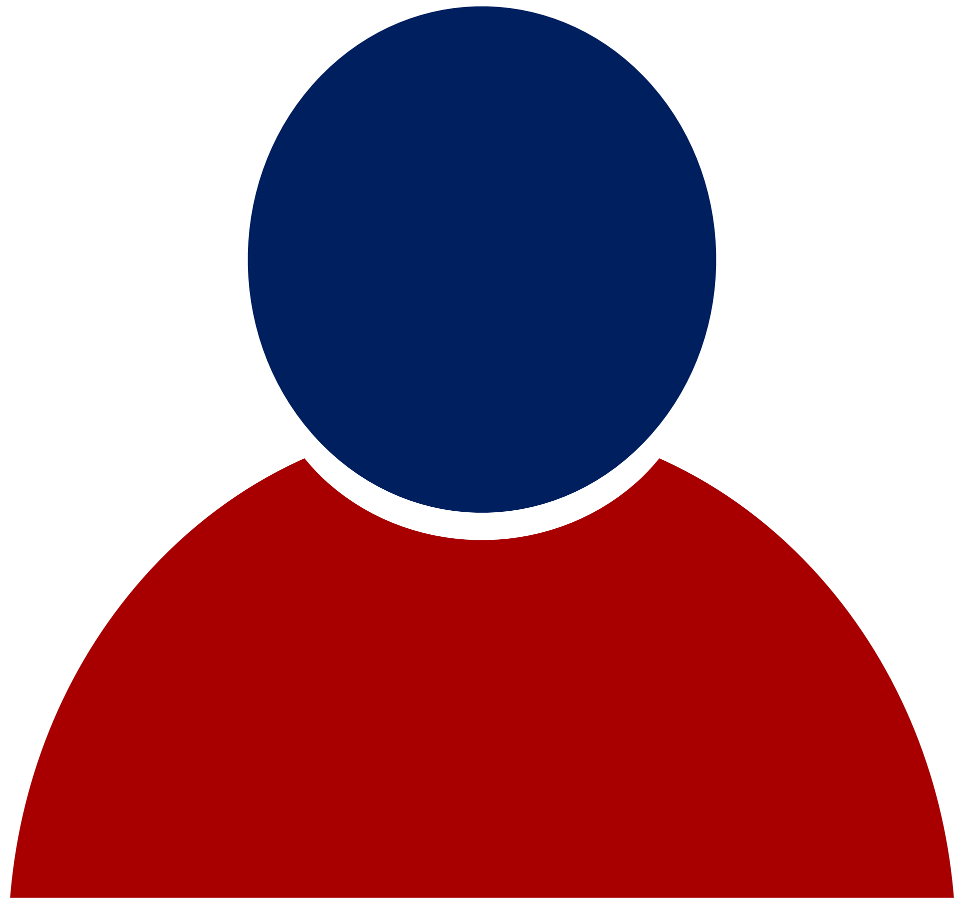
Figure-centric reading patterns with augmentations

In preparation

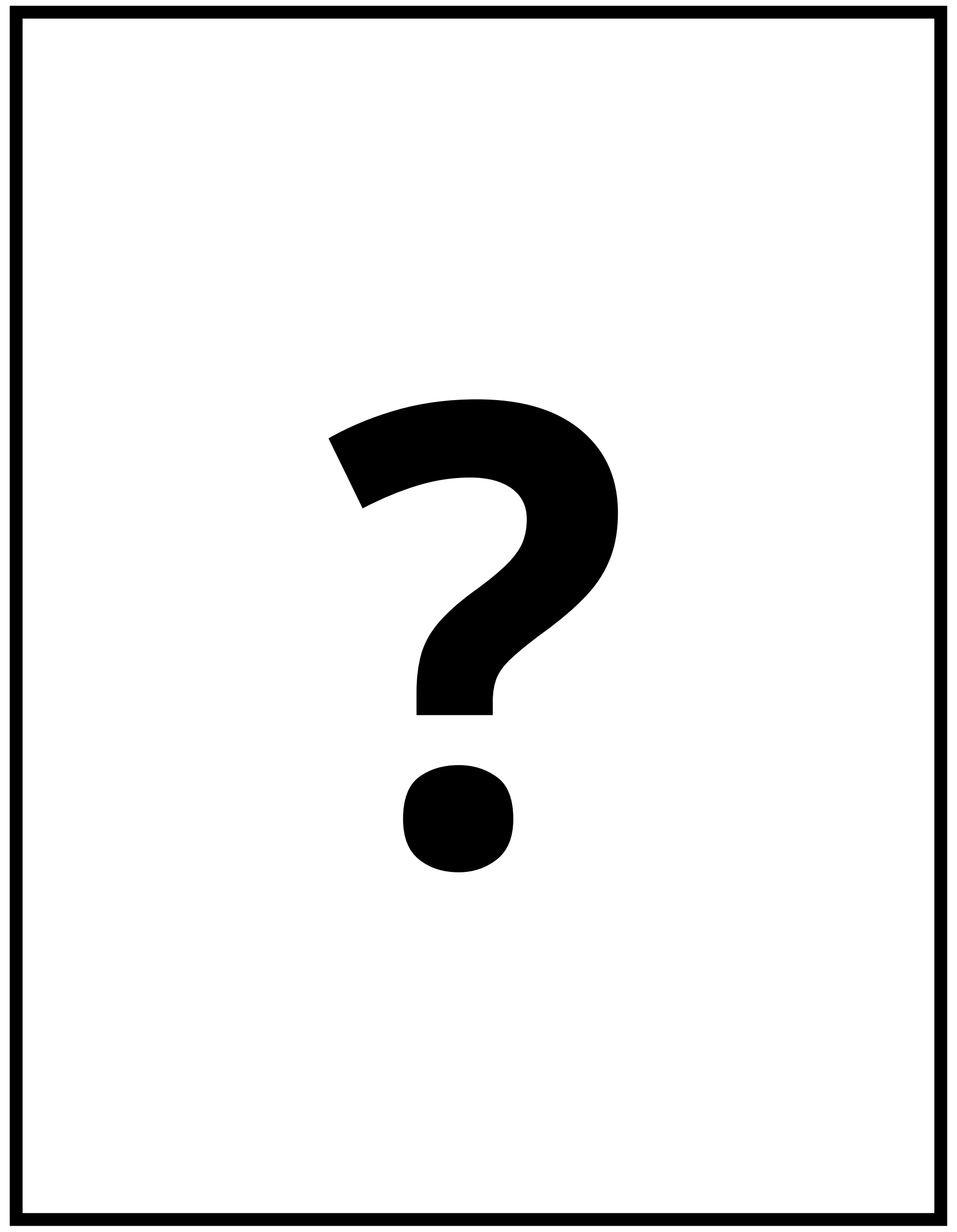
4. Proposed work + timeline

We propose extending the previous study to include a comparison to a baseline.





~30 Penn undergrads



Quiz

Tasks include…

- Revise + resubmit recent paper
- Update prototype
  - Fix errors
  - Regenerate walkthroughs
  - Possibly replace current paper with shorter one
- Write quiz questions
- Run pilot studies (informal practice + feedback sessions)
- Run ~30 user studies (possibly in groups)
- Write and defend dissertation

- **August 4**: proposal presentation

- **September**: begin job search, resubmit paper, prepare prototype
  - Sept. 11: resubmission deadline

- **October**: run user studies

- **November**: address committee feedback, prepare for defense
  - Nov. 24-26: defense (before Thanksgiving)

- **December**: finish dissertation revisions, graduate
  - Dec. 11: last day to deposit dissertation

Thank you! Questions?

- **August 4**: proposal presentation
- **September**: begin job search, resubmit paper, prepare prototype
- **October**: run user studies
- **November**: address committee feedback, prepare for defense
- **December**: finish dissertation revisions, graduate

- Revise + resubmit recent paper
- Update prototype
  - Fix errors
  - Regenerate walkthroughs
  - Possibly replace current paper with shorter one
- Write quiz questions
- Run pilot studies (informal practice + feedback sessions)
- Run ~30 user studies (possibly in