

AI Summaries

Preference Testing Study

May 2025

Jessica Wood

TABLE OF CONTENTS

- 01 Background & Key Findings
- 02 Study Screenshots
- 03 Experiences A, B, & C Findings
- 04 Detailed Findings
- 05 Key Recommendations & Next Steps
- 06 Appendix

Background & Key Findings

AI Summaries Preference Testing Study

Background

Company X is redesigning how search results are displayed, including the addition of an AI-generated summary at the top of the results page. This research aims to identify which elements from the AI summary layout users prefer and the reasons behind their preferences. It also explores how trustworthy users perceive AI summaries to be. Findings from this study will be shared with key stakeholders to help guide the final design decisions.

Research Questions

- Which elements from the designs are preferred and why?
 - *Exp A: AI Summary with bulleted Ref Links and defined Ref section*
 - *Exp B: AI Summary with in-line Ref Links*
 - *Exp C: AI Summary collapsed view with 'Show more' option*
- What types of improvements, if any, do users want to see and why?
- In what cases or for what types of tasks do users rely solely on AI summaries?
- Overall, do users find AI summaries to be reliable?

Method & Users

- Sample N=20
 - 10 AI Developers
 - >15 hours coding/week
 - Familiarity with AI services
 - 10 non-AI developers
- Unmoderated <30-minute sessions conducted via the User Testing platform.
- Users reviewed three design concepts to compare reference layouts and summary display variations (e.g., Show more option).
- User were provided side-by-side comparisons and asked to choose one or no preference.
- Screenshot order was rotated across groups of 10 to minimize order bias.
- Additional insights were gathered through open-ended questions to better understand their preferences and reasoning.

Related Research

- Multiple Accepted Answers Study

Key Findings

- 1. Users expressed a need for more visually prominent reference links (Exp A) and inline citations (Exp B), along with a clearly defined reference section (Exp A)** to improve transparency and make it easier to trace information back to its sources. They suggested design treatments such as bold text, color, or bullets to enhance visual hierarchy within the summary.
- 2. Users preferred summaries to be fully expanded on page load, with the option to collapse,** as this reduced friction and made it easier to quickly scan and assess the content without additional clicks.
- 3. User feedback suggests a desire for a more interactive search results experience within Learn.** Requests to ask follow-up questions or regenerate AI summaries indicate expectations for a dynamic, conversational interface—more aligned with tools like ChatGPT or Copilot than static search results.
- 4. Users' choice between AI tools and traditional search is shaped by both the nature of the task and their comfort with the technology—**AI developers lean on AI for quick, targeted responses, whereas non-AI developers often default to search engines due to routine use and familiarity.
- 5. Users turn to AI summaries within search for quick, straightforward answers,** but shift to traditional search results when a query requires deeper context, multiple perspectives, or critical evaluation.
- 6. Trust in AI summaries depends on source transparency and credibility.** Users emphasize they are more likely to trust—and act on—AI-generated summaries when the information is backed by multiple and reputable sources. Summaries lacking visible or traceable references may lead to skepticism and hesitation.

Experience A Findings

Users found Exp A visually appealing, noting the blue links and clear reference section helped convey credibility and ease of access to sources

- Users appreciated having a dedicated reference section but felt it could be more compact.
- Users were drawn to the blue reference links, with one user suggesting adding a link icon, like a chain, for further clarity.
- The main concern users raised about this layout was that it wasn't clear how the links connected to the summary content.

Recommendations

- Consider designs that visually separate the reference section from the summary to help users clearly distinguish between summary content and supporting sources.
- Consider design treatments like bold text, blue links, or link icons to make reference links more visually distinct and easier for users to spot and engage with.

"I really like this different section of the references because it really helps me to look at the references, so I really like experience A." – P9 AI Dev



Experience B Findings

Users appreciated Exp B linked specific statements to sources via in-line numbers, which made it transparent where the info was coming from and helped the AI summary feel reliable

- Several users appreciated the idea of clickable citation numbers directly in the summary text, enhancing interactivity.

"It tells you that in this AI summary this is where the information is coming from and feels more reliable." – P7 AI Dev

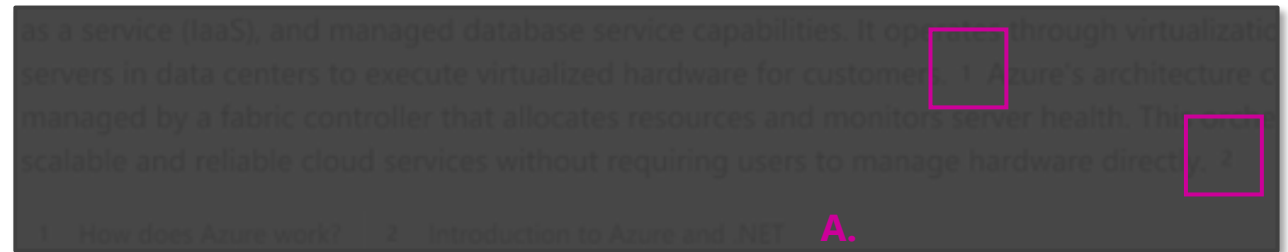
Recommendations

- Consider designs that prioritize clear, inline citation patterns that directly link claims to sources to reinforce transparency and help users feel more confident in the accuracy of AI-generated summaries.

FINDING #3

In Exp B, users noticed the citations but felt they didn't stand out enough, suggesting the need for stronger visual hierarchy to improve clarity and ease of scanning

- While users appreciated the inline citations, many felt they blended in too much with the summary text and suggested making them more visually distinct—such as using bolding or a color design treatment.
- Users found the bottom references [A] more compact, but noted they still needed stronger visual distinction to stand out.
- A few users questioned whether the numbers were clickable as well as the references.



"Within the paragraph, the little parenthetical citations I did not see. So maybe making those a different color, like blue or something like would really make it pop."
– P17 Non-AI Dev

Recommendations

- Consider design treatments like bolding or colors to make the references and inline citations more visually distinct and easier for users to spot and engage with.

"I would make the boxes for both the inline citations and the citations on the bottom like light blue." – P5 AI Dev

Experience C Findings

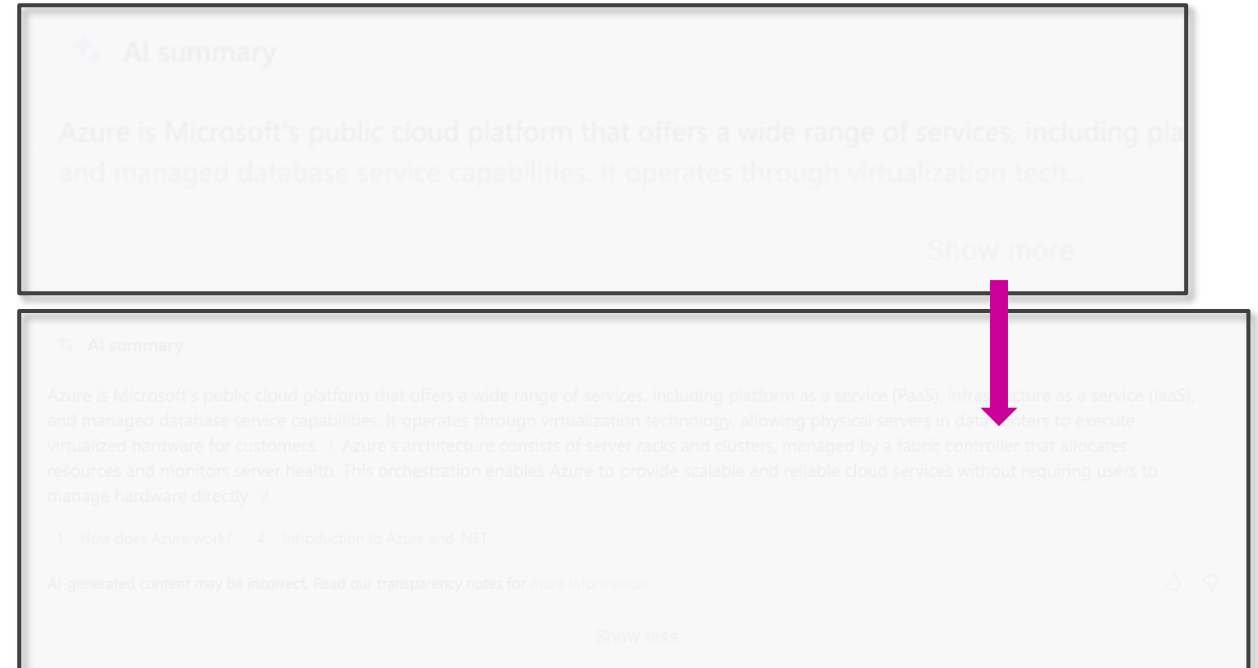
FINDING #4

Experience C was polarizing—while some appreciated the flexibility for longer content, the majority found the default hidden state added friction and made it difficult to scan quickly

- Some users appreciated the ability to collapse or expand the summary when appropriate, especially for longer content. This gave them a sense of control over their reading experience.
- Other users felt the 'Show more' option didn't provide enough context and suggested the design should better indicate the summary is longer and will expand significantly.

Recommendations

- Explore designs that default the AI summary to an expanded view with the option to collapse, allowing users to easily access the full content while retaining control over their viewing experience.



Detailed Findings

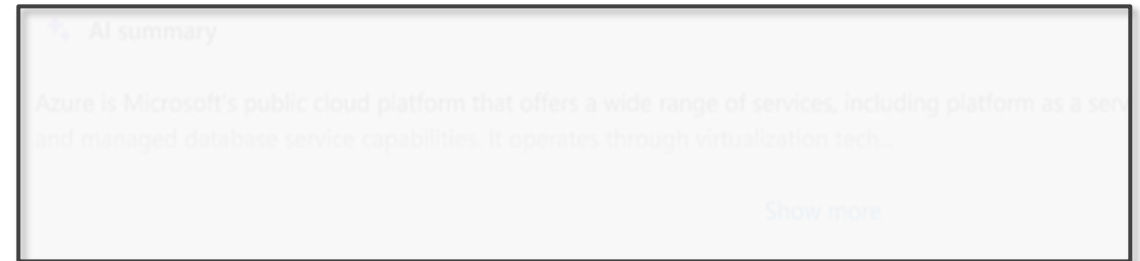
Users preferred summaries to load fully expanded, with the option to collapse, as it reduced friction by letting them decide when and how to engage with the content

- Users acknowledged AI summaries could be longer in a real environment and expressed a need for a collapse option—but still preferred them open by default. This preference also appeared in Multiple Accepted Answers study.
- Some users noted that if a “Show more” link is used, enough text should be shown upfront to signal that additional content is worth expanding.
- There was no clear preference for the wording—terms like “Show more” and “Hide” received mixed reactions, with no consensus on which was clearer.

Recommendations

- Repeat: Explore designs that default the AI summary to an expanded view with the option to collapse, allowing users to easily access the full content while retaining control over their viewing experience.
- Consider researching which terminology is used across Company X products for collapse/expand functionality and running a quick competitive scan to see how other platforms label similar interactions.
- Should it be decided to pursue a design in which the AI summary appears collapsed at default, explore ways to preview more of the content to help users decide whether to expand and continue reading.

Experience C



"I prefer to see everything by default, but I do like having the option to shorten it, to expand it or collapse it." – P10 AI Dev



Users expressed that combining B's inline citations with A's separate reference section would improve transparency by making it easier to quickly trace information back to its sources

- When choosing between Experiences A and B, users consistently favored a hybrid approach—combining A's separate reference section and visually distinct links with B's inline citation style for better clarity and traceability.

"I prefer Experience A because it does better job of explaining as references, but I like that B has the numbering like one and two and it shows exactly where one and two is pulled from. I think it has to be a combination of both of these." – P1 AI Dev



"I prefer experience A in terms of the references, but one thing is it doesn't talk about which line from which reference it was taken. From that perspective, maybe B would be the one" – P14 Non-AI Dev



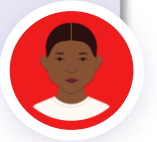
Recommendations

- Consider exploring a hybrid design that integrates key elements from both Experience A and Experience B to improve clarity and traceability, helping users more easily understand and trust the AI summary.

Preference between AI tools and traditional search depends on task and familiarity—AI devs favor AI for fast, contextual answers, while non-AI devs stick with search engines out of habit

- AI developer users shared that tools like ChatGPT and Copilot save them time by eliminating the need to sift through multiple websites for answers.
- Non-AI developers reported they tended to use traditional search engines out of habit and believed they offered more relevant local or regional information than AI tools like ChatGPT or Copilot.
- This difference in perspective may stem from AI developers' familiarity with the technology—they work closely with AI tools and better understand their capabilities, while non-developers may be less aware of the full functionality and potential of these tools.

"The instinct is still there to automatically go to Google if I'm in need of information" – P12 Non-AI Dev



"I use Google because that's what I'm used to and it works just fine for me." – P20 Non-AI Dev



"When I need quick information, I tend to use AI tools because it actually gives you the specific answer you need for the question. However, if you go through, Bing or Google search, you'd actually have to filter through to get the specific answer you need. But when you use AI tools, you get to nail the answer on the head." – P2 AI Dev



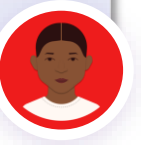
A few users expressed interest in following up with additional questions after reading the AI summary, suggesting a preference for a more iterative, dialog-based search experience

- A couple of users vocalized the need to ask follow-up questions about the AI summary, indicating a desire for the interaction to function more like ChatGPT or Copilot.
- A few users wanted the ability to regenerate a response or refine their question, suggesting a preference for a more interactive, ChatGPT or Copilot-like experience. A similar finding was reported in the Prompt and Training study.

Recommendations

- Consider conducting exploratory research to better understand how users interact with Company X search results, including their behaviors, expectations, and mental models.

"Sometimes I have to reformulate the prompt because I had this question that was very open and I want just want specific details about it. So I have to rewrite the prompt and say, I want this answer." – P11 Non-AI Dev



"It would be really nice if I was able to like ask maybe one or two follow up questions. I know that generating AI text is really expensive, but for me personally, I've had situations where the AI summary gives a little bit of information, but it doesn't really answer my question fully." – P5 AI Dev



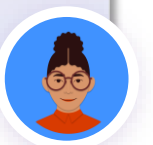
When performing a Google/Bing search, users rely on the AI summary for simple, fact-based questions, but prefer the traditional search results for complex, nuanced, or debatable topics

- Users trust AI summaries for objectively true facts—like “How tall is Mt. Everest?” or “How do I boil an egg?”—where answers are straightforward and widely agreed upon.
- Users expressed that they’ll use search engines when they know there are multiple valid perspectives, ambiguous context, or if their search requires detailed exploration.

“If it is questions around a theory, logic, math then I will typically go with the AI generated summary.” – *P18 Non-AI Dev*



“if I’m looking for direct answer, let’s say what’s the height of Mount Everest then I look at the AI summary cause it will give you the direct answer immediately.” – *P1 AI Dev*



Trust in AI summaries is closely tied to the perceived reliability of the sources they reference, suggesting clear, traceable citations should be a core component of any Learn AI summary

- Users stressed that without clear, connected references, they're unlikely to trust or act on AI-generated summaries. Credible sources are key to validating the information.
- Users expressed they tend to trust AI summaries that include multiple reputable sources—whereas summaries with no or only one source may lead to skepticism.
- A few users specifically said they would dismiss AI summaries that cite non-reputable sources, such as Reddit, viewing them as unreliable.

Recommendations

- Consider exploring design solutions that integrate references directly into the AI summary to improve traceability and enhance the perceived reliability of the information presented.

"I do think that they [AI summaries] should be kind of stemming from reliable sources. So, if the link comes from like an .edu post or from a peer reviewed source, those sources will make it more reliable" – P4 AI Dev



"I trust at least five references so that I can fully trust the claims of the AI summary and so that I feel that it's not really cherry picking anything just to fit in the narrative." – P3 AI Dev



"[To make AI summaries more reliable] I think perhaps if it cited the sources in which it's gathering this information from, are some of the more popular or more credible places in which parts of this information were taken from" – P17 Non-AI Dev



Key Recommendations & Next Steps

Key Recommendations

1. **Consider designs where reference links are directly connected to specific text** or claims within the AI summary to enhance clarity and transparency.
2. **Consider designs that visually separate the reference section from the summary** to help users clearly distinguish between summary content and supporting sources.
3. **Explore design treatments for reference links/citations that are visually distinct and easy to identify** within the summary, enabling users to quickly recognize and interact with supporting sources.
4. **Explore designs that default the AI summary to an expanded view with the option to collapse**, allowing users to easily access the full content while retaining control over their viewing experience.

Next Steps

The following steps are research suggestions and are not placed in order of priority. Priority of work to be determined by product partners, resources, and research + design.

1. **Consider a follow-up moderated study to explore** whether users expect a more conversational, chat-like search experience rather than static results when on Learn.
2. **Consider reviewing Company X products and conducting a competitive analysis** to identify how collapse and expand functionality is labeled, helping inform the most effective wording.
3. **Consider running an unmoderated click test to understand whether users notice, prioritize, and comprehend** key elements based on their visual styling (e.g., size, color, placement, contrast).
4. **Consider conducting a concept test to explore different design treatments for expanded vs. collapsed summaries to help assess which approach best supports user preferences** for low-friction engagement and influences their willingness to engage with the summary content.

THANK YOU

Appendix:

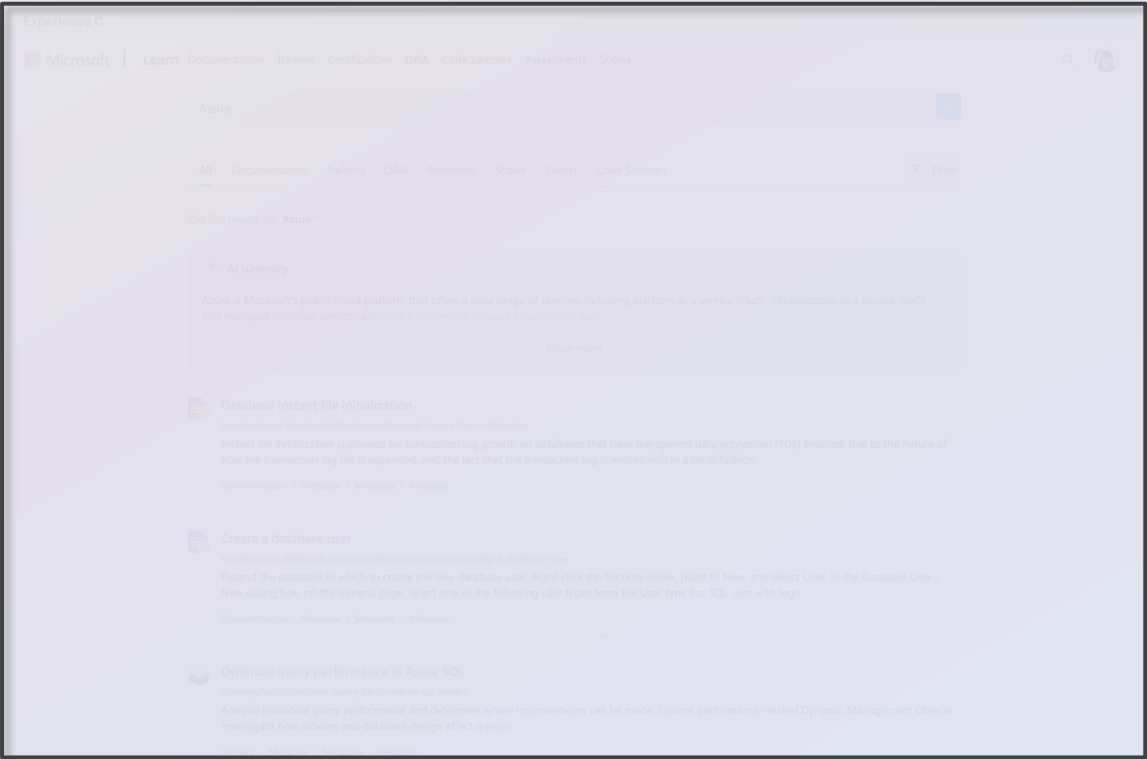
While B’s reference layout was preferred over A’s, non-AI developers leaned toward A—revealing a preference split based on user background

All Users		AI Devs		Non-AI Devs	
Prefers A	Prefers B	Prefers A	Prefers B	Prefers A	Prefers B
8	12	2	8	6	4



Both user groups preferred the expanded summary view (B) over the collapsed version (C), as it let them see the full context without needing to click to reveal more

All Users		AI Devs		Non-AI Devs	
Prefers B	Prefers C	Prefers B	Prefers C	Prefers B	Prefers C
14	6	7	3	7	3



Non-AI devs and AI devs agreed the information in AI generated summaries to be reliable

Users were asked their agreement to the following statement: *The information in AI-generated summaries within search results is reliable.*

N=20

