Behavioral and Cognitive Data

Insights for User-Centric Web Design

HTI Endterm

16/12/2024

Anushka Desai, Devesh Shah, Shaurya Sighadia

Problem Statement

How can user-centric web design improve **content clarity, navigation, and user experience** by analyzing **behavioral and cognitive data** ?

Problem:

- Current web interfaces often confuse users due to unclear content organization and navigation.
- Design decisions often rely on guesswork instead of data-driven insights.
- This leads to frustration, inefficient information retrieval, and low user satisfaction.

Potential Impact:

- **Higher User Satisfaction:** Users find what they need faster, with less frustration.
- **Increased Efficiency:** Streamlined interactions reduce time and effort.
- **Competitive Advantage:** Data-driven, user-centric design fosters trust, loyalty, and long-term innovation.

Potential Applications:

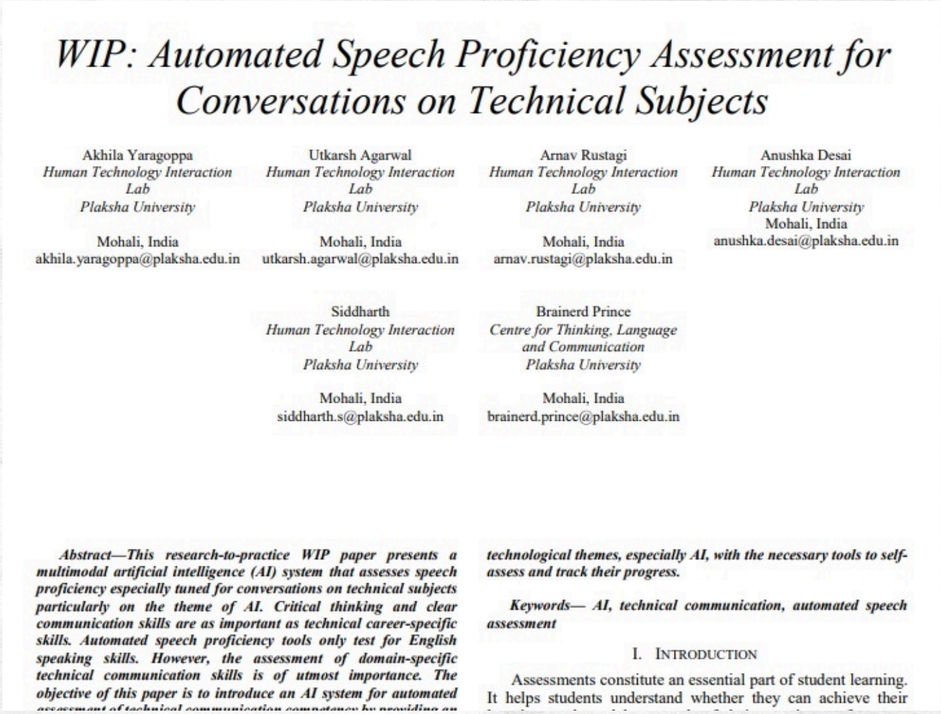
- **Enhanced Usability:** Create intuitive navigation and layouts using real-time behavioral data.
- **Clearer Content Delivery:** Identify and fix points where users struggle to understand information.
- **Continuous Design Optimization:** Use cognitive and behavioral metrics for ongoing, evidence-based design improvements.

Literature Survey

1. Technical Divide in Speech Proficiency

- **Tech:** demonstrate **higher speech proficiency** when discussing topics within their area of expertise, such as AI or machine learning (ML). They are **more comfortable discussing technical subjects**, and their responses tend to be more relevant and aligned with the topic.

- **Non Tech:** show **lower proficiency** when asked to discuss unfamiliar technical topics. Their speech contains more **disfluencies when they speak about AI or machine learning**, as these subjects are outside their expertise.



WIP: Automated Speech Proficiency Assessment for Conversations on Technical Subjects

Akhila Yaragoppa
Human Technology Interaction Lab
Plaksha University
Mohali, India
akhila.yaragoppa@plaksha.edu.in

Utkarsh Agarwal
Human Technology Interaction Lab
Plaksha University
Mohali, India
utkarsh.agarwal@plaksha.edu.in

Arnav Rustagi
Human Technology Interaction Lab
Plaksha University
Mohali, India
arnav.rustagi@plaksha.edu.in

Anushka Desai
Human Technology Interaction Lab
Plaksha University
Mohali, India
anushka.desai@plaksha.edu.in

Siddharth
Human Technology Interaction Lab
Plaksha University
Mohali, India
siddharth.s@plaksha.edu.in

Brainerd Prince
Centre for Thinking, Language and Communication
Plaksha University
Mohali, India
brainerd.prince@plaksha.edu.in

***Abstract*—This research-to-practice WIP paper presents a multimodal artificial intelligence (AI) system that assesses speech proficiency especially tuned for conversations on technical subjects particularly on the theme of AI. Critical thinking and clear communication skills are as important as technical career-specific skills. Automated speech proficiency tools only test for English speaking skills. However, the assessment of domain-specific technical communication skills is of utmost importance. The objective of this paper is to introduce an AI system for automated**

technological themes, especially AI, with the necessary tools to self-assess and track their progress.

***Keywords*— AI, technical communication, automated speech assessment**

I. Introduction

Assessments constitute an essential part of student learning. It helps students understand whether they can achieve their

Plaksha HTI Lab: WIP - Automated Speech Proficiency Assessment for Conversations on Technical Subjects (2024)

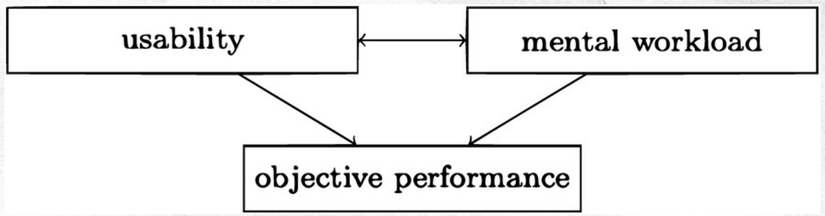
2. What is Cognitive Load?

- In the context of UX, it measures **how much information a user must process** to **interact with a system effectively.**
- Types of Cognitive Load:
  - Intrinsic: complexity of the **task itself.**
  - Extraneous: unnecessary mental effort caused by **poor designs.**
  - Germane: **mental effort** directed toward **constructing meaningful connections**
- A well-designed interface **minimizes extraneous load** while **optimizing germane load**, allowing users to focus on the task at hand without unnecessary distractions.

Interaction Design Foundation - IxDF. (2016, June 5). What is Cognitive Load?. Interaction Design Foundation - IxDF. https://www.interaction-design.org/literature/topics/cognitive-load

3. Experienced mental workload, perception of usability, their interaction and impact on task performance

- **Inverse Relationship:** Higher usability reduces mental workload, improving task performance and user satisfaction; poor usability increases cognitive effort and frustration.
- **Task Familiarity & Feedback:** Familiar users experience lower workload, while real-time system feedback enhances usability perception and reduces cognitive effort.
- **Usability > Performance:** Perceived usability is a stronger predictor of user satisfaction than task performance, even when tasks are successfully completed.



1 Materials and methods

Widely employed in the broader field of HCI, usability and mental workload are two constructs from the discipline of Human Factor, with no limpid and broad definitions. Since their inception, there has been an intense debate about their assessment and measurement [19–21]. Although multidimensional and complex, their usefulness for describing the user experience and informing interaction, interface and system design is beyond doubt.

1.1 Usability

The amount of literature covering definitions of usability, its frameworks and methodologies is significant [22–26]. An early definition by Eason [27] was 'the degree to which users are able to use the system with the skills, knowledge, stereotypes and experience they can bring to bear'. This definition is mainly connected to the ease-of-use, however, there are more factors influencing usability. For example, a definition provided by the ISO (International Organisation for Standardisation), is 'The extent to which a product can be used by specified users to

Longo L (2018) Experienced mental workload, perception of usability, their interaction and impact on task performance. PLoS ONE 13(8): e0199661. https://doi.org/10.1371/journal.pone.0199661

4. Pupil dilation as an index of effort in cognitive control tasks: A review

- **Effort Tracking:** Pupil dilation reliably reflects cognitive effort, increasing with task complexity, especially in working memory, decision-making, and multitasking tasks.
- **Dynamic Feedback:** Dilation peaks during the most demanding task phases, offering real-time feedback on mental load and task difficulty.
- **Context Sensitivity:** Environmental factors (e.g., light, fatigue, emotions) influence dilation, requiring careful control for accurate measurement.

Pupil dilation as an index of effort in cognitive control tasks: A review

Pauline van der Wel [1] · Henk van Steenbergen [1,2]

Published online: 12 February 2018

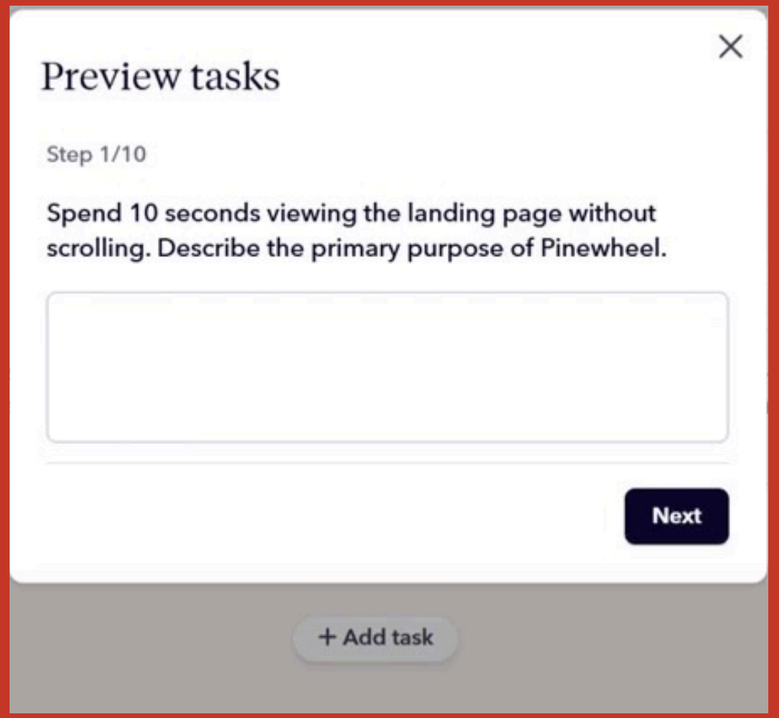
Abstract
Pupillometry research has experienced an enormous revival in the last two decades. Here we briefly review the surge of recent studies on task-evoked pupil dilation in the context of cognitive control tasks with the primary aim being to evaluate the feasibility of using pupil dilation as an index of effort exertion, rather than task demand or difficulty. Our review shows that across the three cognitive control domains of updating, switching, and inhibition, increases in task demands typically leads to increases in pupil dilation. Studies show a diverging pattern with respect to the relationship between pupil dilation and performance and we show how an effort account of pupil dilation can provide an explanation of these findings. We also discuss future directions to further corroborate this account in the context of recent theories on cognitive control and effort and their potential neurobiological substrates.

Keywords Pupil · Effort · Attention · Cognitive control · Inhibition · Switching · Updating · Behavioral performance · Resource allocation · Cognitive load

van der Wel, P., van Steenbergen, H. Pupil dilation as an index of effort in cognitive control tasks: A review. Psychon Bull Rev 25, 2005–2015 (2018). https://doi.org/10.3758/s13423-018-1432-y

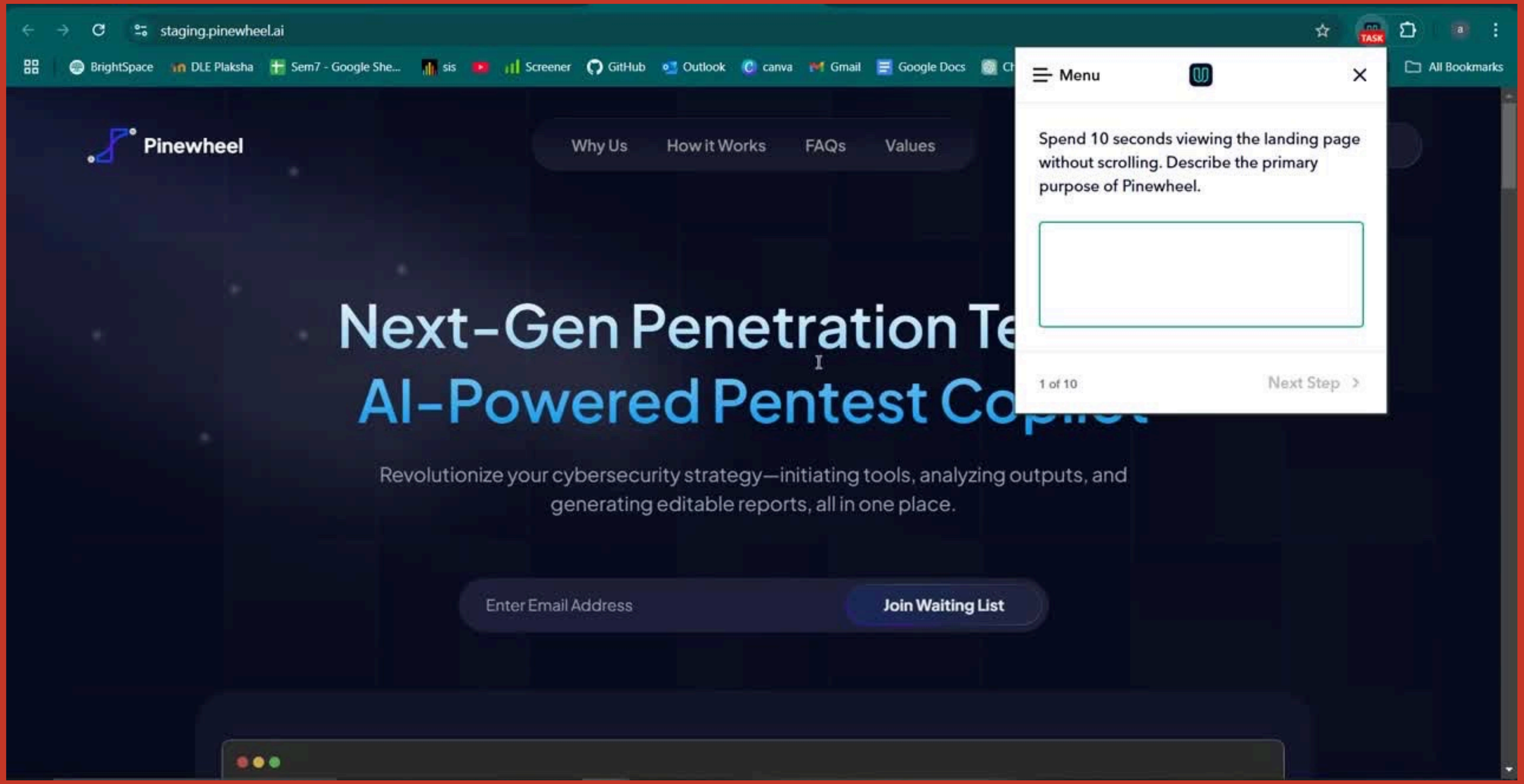
Experiment Protocol

30 participants



9 tasks

Experiment SetUp


staging.pinewheel.ai
BrightSpace
DLE Plaksha
Sem7 - Google She...
sis
Screener
GitHub
Outlook
canva
Gmail
Google Docs
All Bookmarks
Pinewheel
Why Us
How it Works
FAQs
Values
Next-Gen Penetration Te
AI-Powered Pentest Co
Revolutionize your cybersecurity strategy—initiating tools, analyzing outputs, and generating editable reports, all in one place.
Enter Email Address
Join Waiting List
Menu
Spend 10 seconds viewing the landing page without scrolling. Describe the primary purpose of Pinewheel.
1 of 10
Next Step

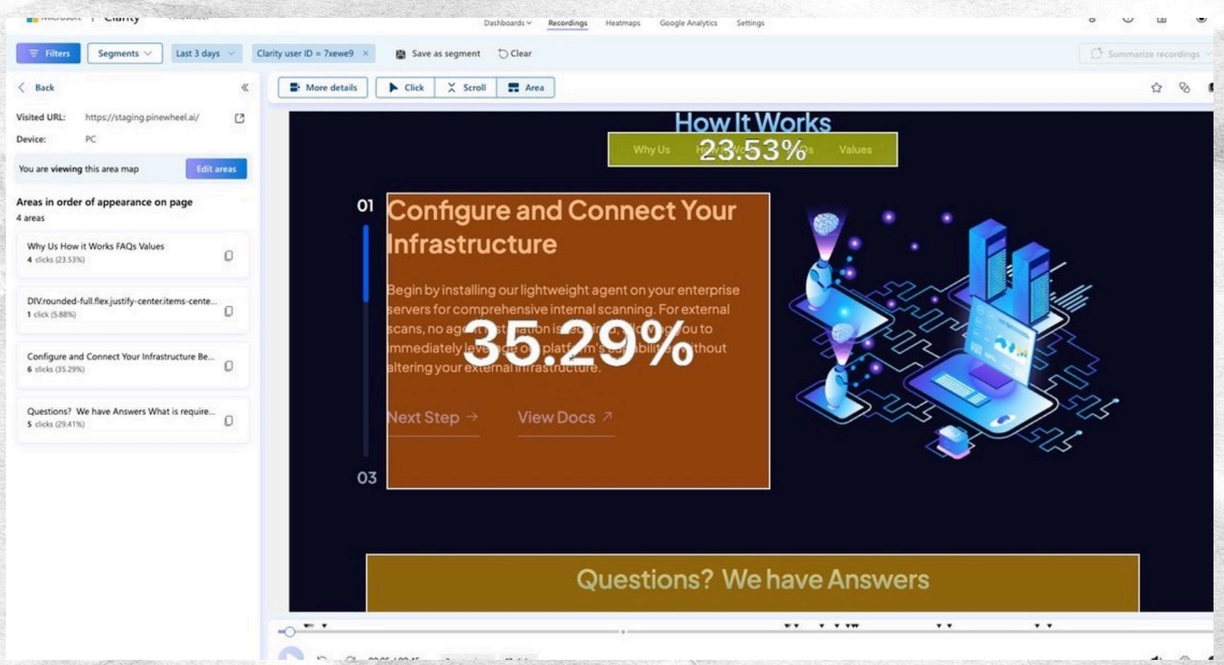
Experiment Apparatus

- **Microsoft Clarity** for website interaction data
- **UserBrain** for live survey data
- **IR Camera + Rasberry Pi + 3D Printed Components** for video data
- **mEye** for pupil data extraction from videos

Data Collected

Website Interaction

- **Section Interaction Data (Clicks and Area):**
  - Hero Section
  - Navigation Bar
  - How It Works Section
  - Features Section
  - FAQs Section
- **Dead Clicks:** These are clicks on non-interactive or unresponsive areas, highlighting design flaws or misleading elements that confuse users.
- **Back Clicks:** Represents user actions to return to a previous page or section, indicating potential navigation issues or dissatisfaction with the current content.



Tasks

🧩 Understanding the Platform's Purpose

- Task 1 evaluates how effectively the landing page communicates Pinewheel's core purpose at **first glance.**
- Task 9 collects **suggestions,** helping identify gaps that hinder **content absorption.**

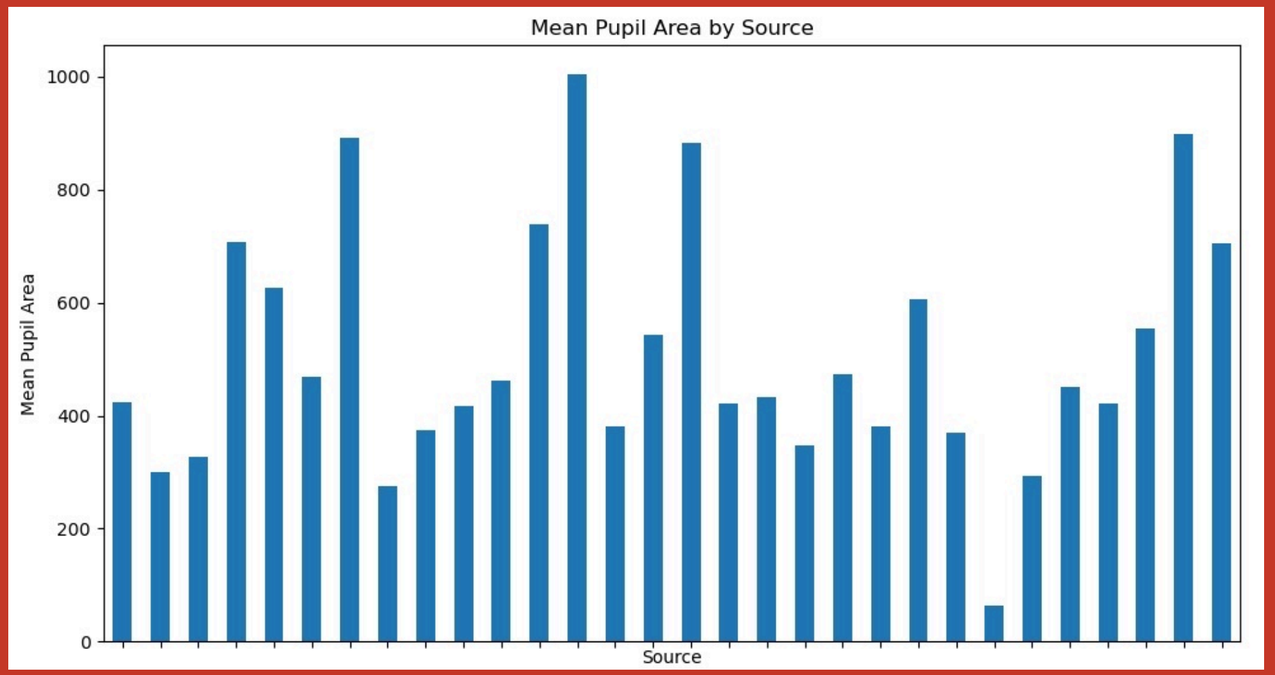
🔍 Assessing Content Clarity and Tools

- Task 2 focuses on understanding Pinewheel’s approach to **data confidentiality**
- Tasks 3, 4, and 5 assess how **well the platform communicates its tools, onboarding process, and feature benefits**
- Navigate and identify key functionalities.

🚀 Navigability and User Impressions

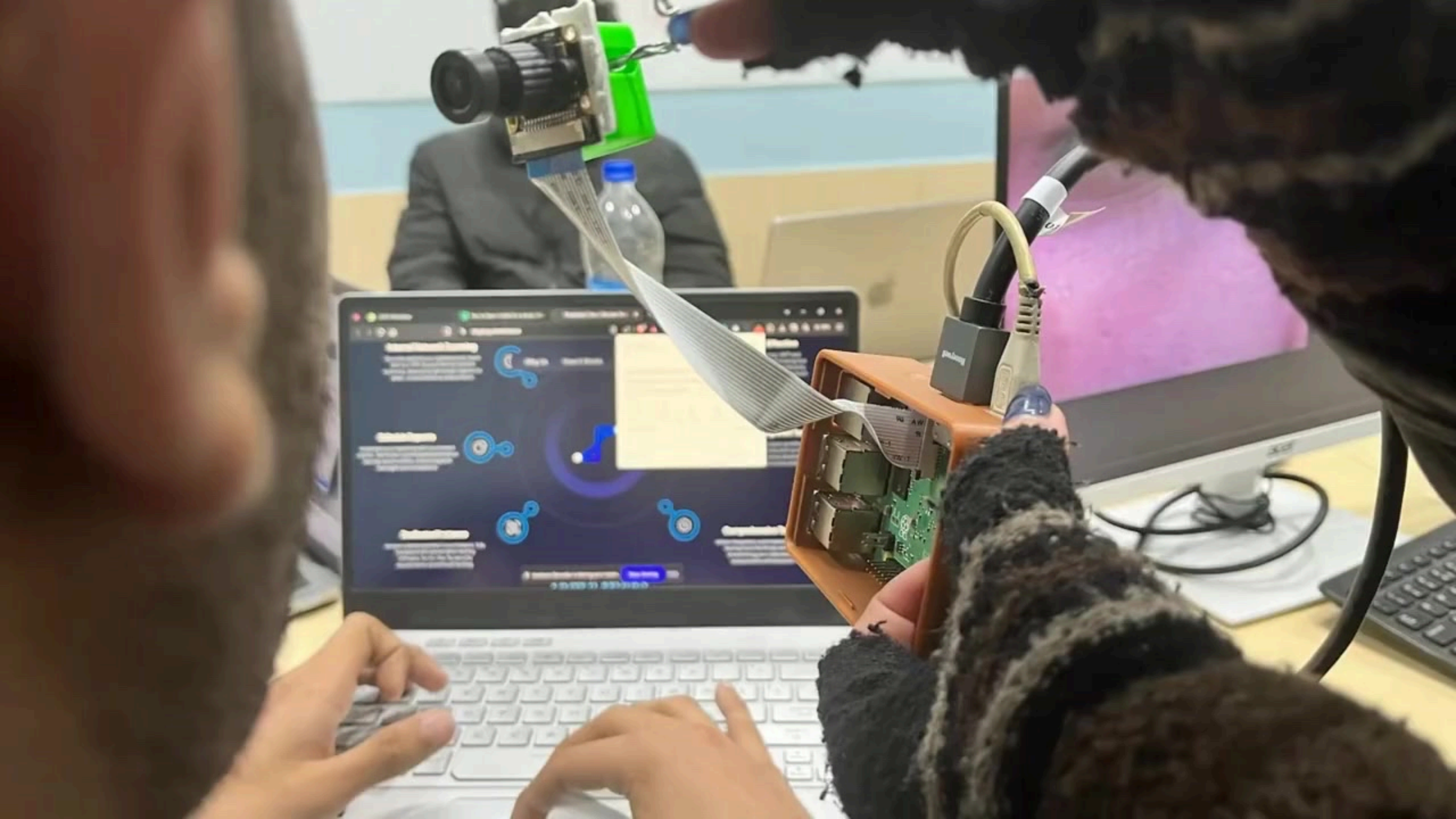
- Task 6 measures how easy it is for users to **navigate** Pinewheel's website.
- Tasks 7 and 8 evaluate **user satisfaction and decision-making factors.**

Pupil Data


Mean Pupil Area by Source
Mean Pupil Area
1000
800
600
400
200
0
Source

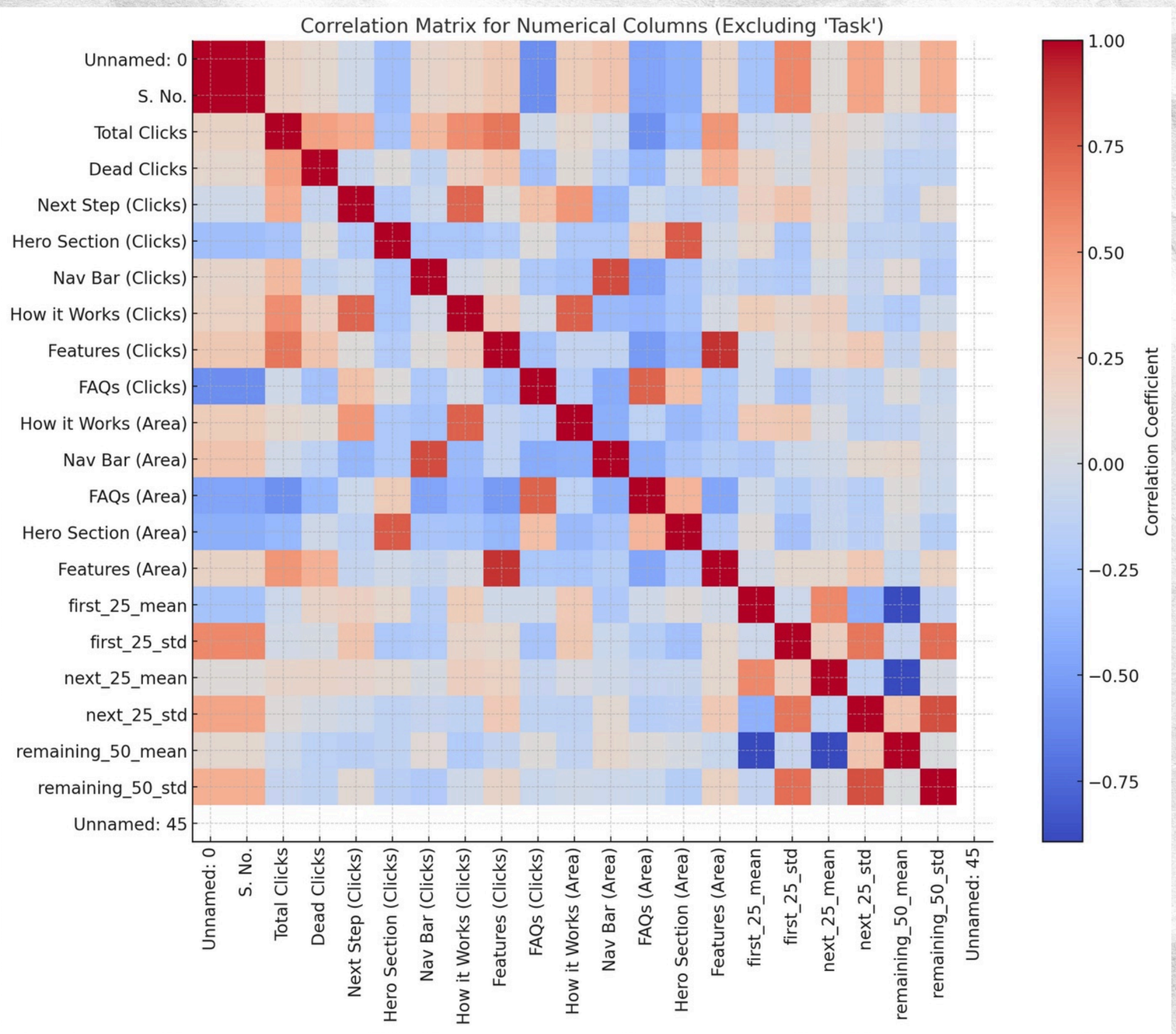
Ethical Considerations

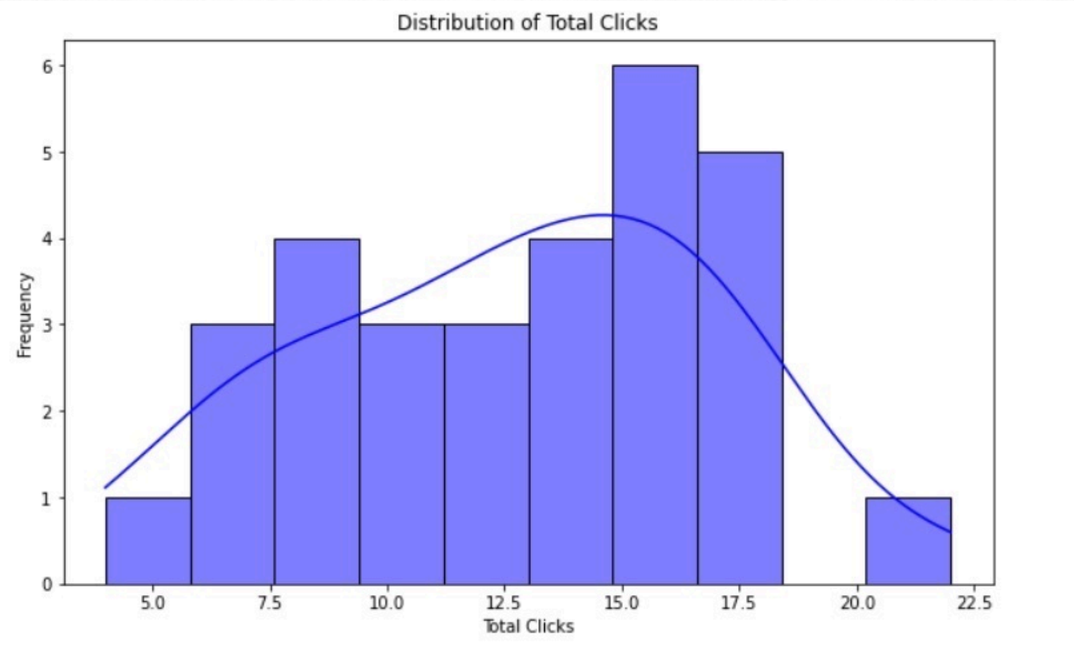
- Participants consented to their data being used for the study.
- Participants were anonymized post survey.
- Participants were allowed to pause and withdraw at any time.
- Transparency & Debriefing.

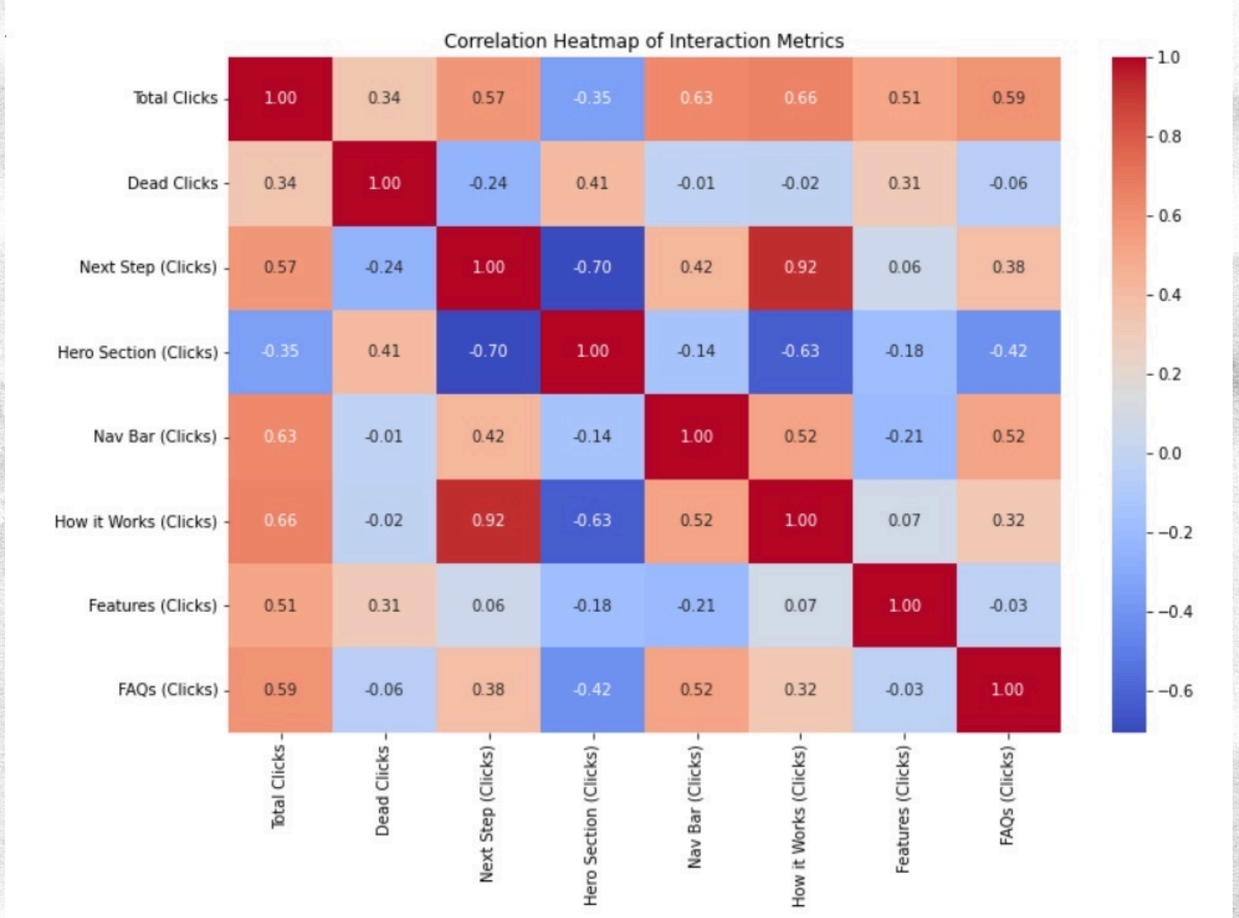


Analysis

1


Correlation Matrix for Numerical Columns (Excluding 'Task')
Unnamed: 0
S. No.
Total Clicks
Dead Clicks
Next Step (Clicks)
Hero Section (Clicks)
Nav Bar (Clicks)
How it Works (Clicks)
Features (Clicks)
FAQs (Clicks)
How it Works (Area)
Nav Bar (Area)
FAQs (Area)
Hero Section (Area)
Features (Area)
first_25_mean
first_25_std
next_25_mean
next_25_std
remaining_50_mean
remaining_50_std
Unnamed: 45
Correlation Coefficient
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75


Distribution of Total Clicks
Frequency
Total Clicks
0
1
2
3
4
5
6
5.0
7.5
10.0
12.5
15.0
17.5
20.0
22.5


Correlation Heatmap of Interaction Metrics
Total Clicks
Dead Clicks
Next Step (Clicks)
Hero Section (Clicks)
Nav Bar (Clicks)
How it Works (Clicks)
Features (Clicks)
FAQs (Clicks)
1.00 0.34 0.57 -0.35 0.63 0.66 0.51 0.59
0.34 1.00 -0.24 0.41 -0.01 -0.02 0.31 -0.06
0.57 -0.24 1.00 -0.70 0.42 0.92 0.06 0.38
-0.35 0.41 -0.70 1.00 -0.14 -0.63 -0.18 -0.42
0.63 -0.01 0.42 -0.14 1.00 0.52 -0.21 0.52
0.66 -0.02 0.92 -0.63 0.52 1.00 0.07 0.32
0.51 0.31 0.06 -0.18 -0.21 0.07 1.00 -0.03
0.59 -0.06 0.38 -0.42 0.52 0.32 -0.03 1.00
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6

2

```
{'Theme 1': ['automate',
  'pinewheel',
  'ai',
  'testing',
  'security',
  'features',
  'using',
  'platform',
  'offered',
  'penetration'],
 'Theme 2': ['features',
  'cybersecurity',
  'platform',
  'testing',
  'penetration',
  'like',
  'tasks',
  'page',
  'gobuster',
  'wpscan'],
 'Theme 3': ['security',
  'ai',
  'video',
  'tool',
  'focus',
  'demo',
  'efficiency',
  'insights',
  'influenced',
  'powered']}
```

We analyzed the open-ended responses from Tasks 1, 8, and 9 to uncover key themes and insights. These tasks were designed to test the following:

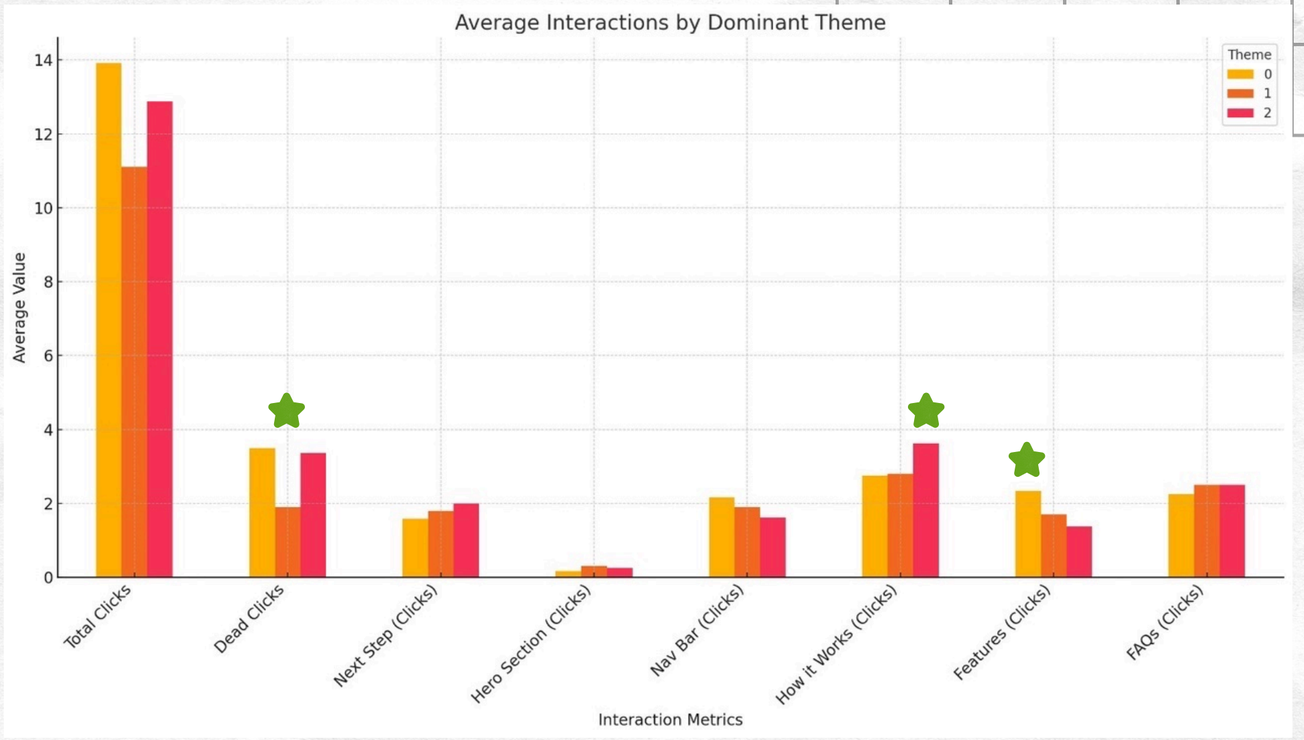
- Task 1: Understand users' first impressions of Pinewheel after briefly viewing the landing page.
- Task 8: Identify factors influencing users' decisions, such as features, usability, or design.
- Task 9: Gather suggestions for improvement, focusing on unmet expectations or lingering questions.

By processing and analyzing these responses using **topic modeling.**

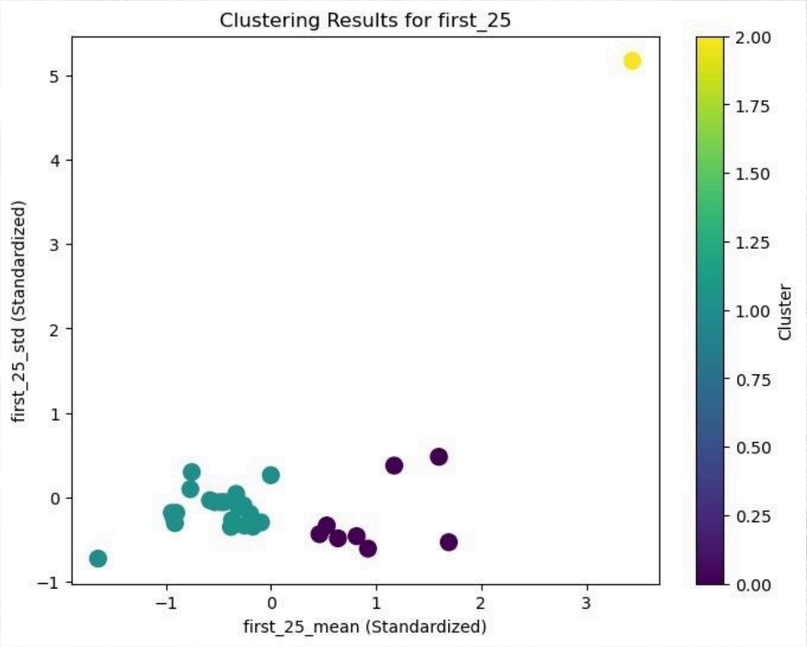
Theme 1: Automation and Penetration Testing:
Keywords: **automate, pinewheel, ai, testing, security, features, platform, penetration.**
Focus: Pinewheel's role in automating security testing processes with AI-driven features.

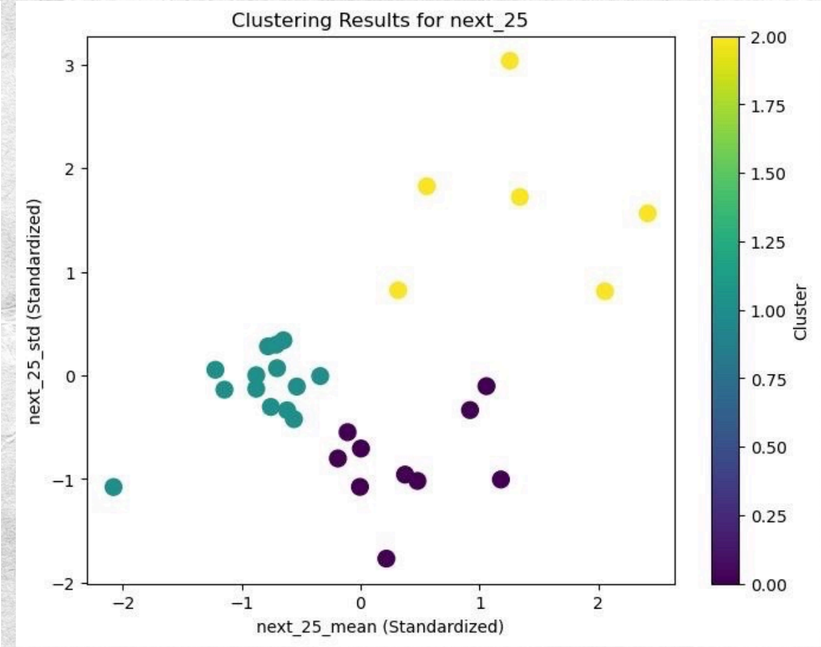
Theme 2: Cybersecurity Features and Tools:
Keywords: **features, cybersecurity, platform, penetration, gobuster, wpscan, tasks, page.**
Focus: Users' engagement with specific cybersecurity features and tools offered on the platform.

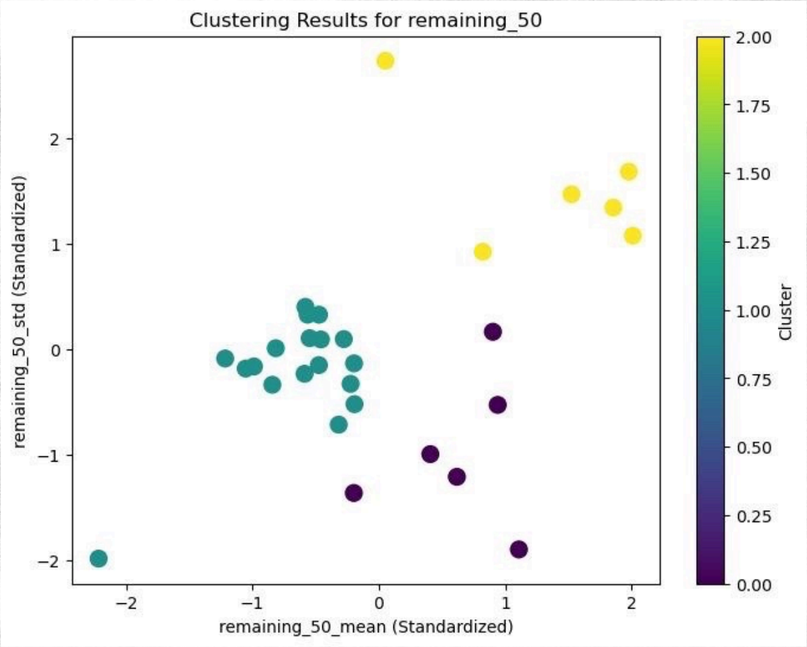
Theme 3: AI Efficiency and Demonstration:
Keywords: **security, ai, video, tool, demo, efficiency, focus, insights.**
Focus: Users value efficiency, insights, and suggest visual demonstrations (e.g., videos) for better understanding.

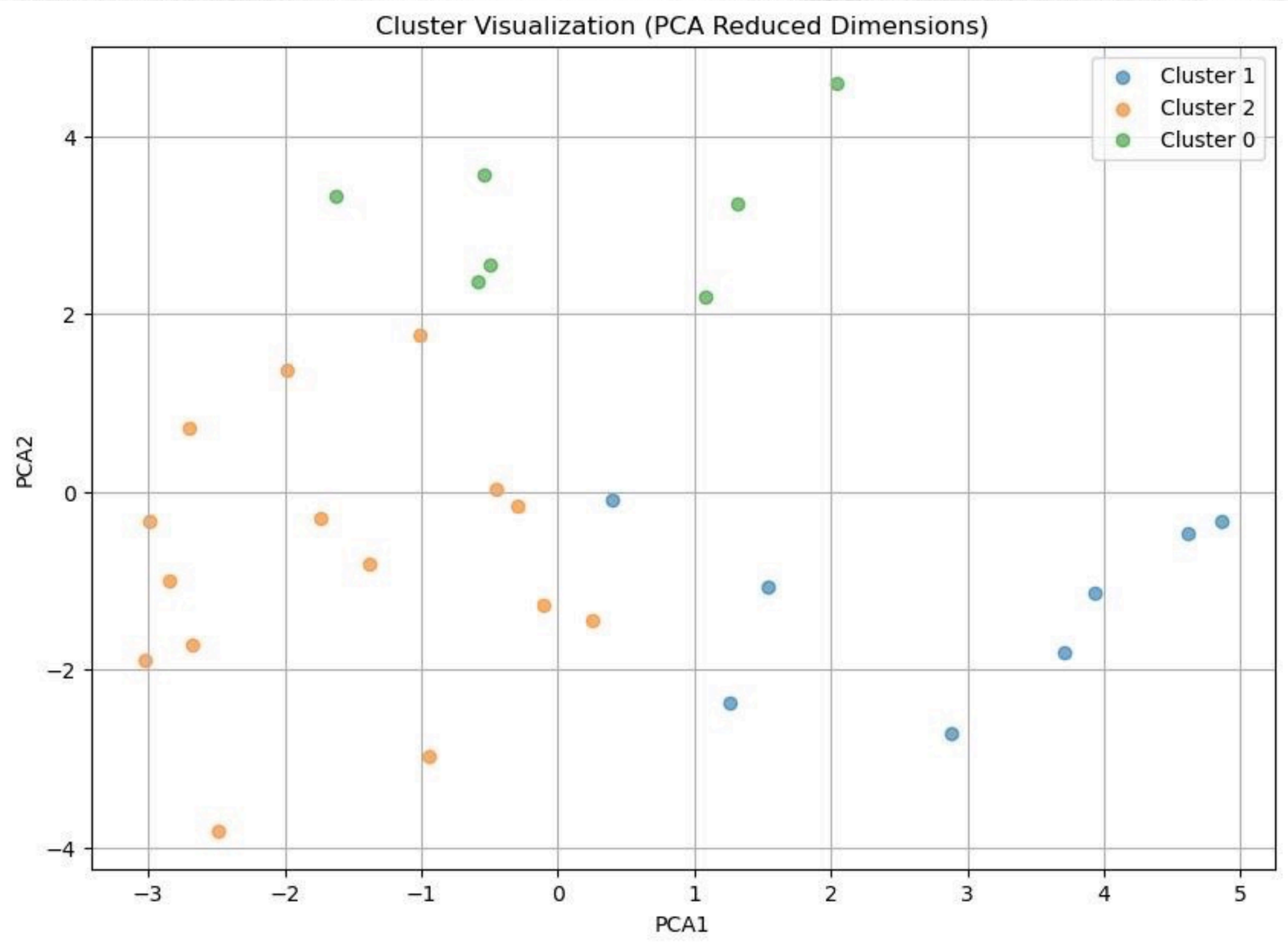


3


Clustering Results for first_25
first_25_std (Standardized)
first_25_mean (Standardized)
Cluster


Clustering Results for next_25
next_25_std (Standardized)
next_25_mean (Standardized)
Cluster


Clustering Results for remaining_50
remaining_50_std (Standardized)
remaining_50_mean (Standardized)
Cluster


Cluster Visualization (PCA Reduced Dimensions)
Cluster 1
Cluster 2
Cluster 0
PCA2
PCA1
4
2
0
−2
−4
−3
−2
−1
0
1
2
3
4
5

Cluster 0:

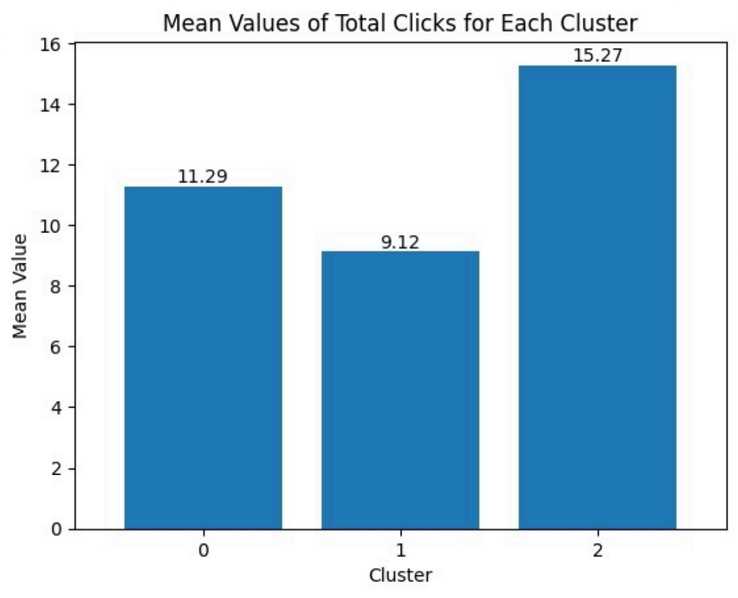
- High confusion initially (negative pupil area early on)
- Difficulty in navigation (high total and dead clicks)
- Successful with Task 2 but struggled significantly with Task 3
- Found instructions unclear, needing prior technical knowledge
- Ended up using technical tools and terms in their feedback, indicating technical users.

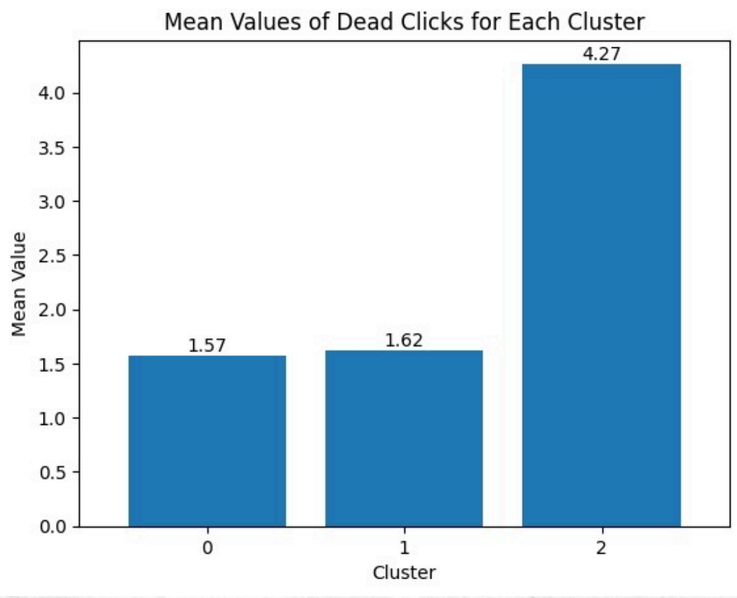
Cluster 1:

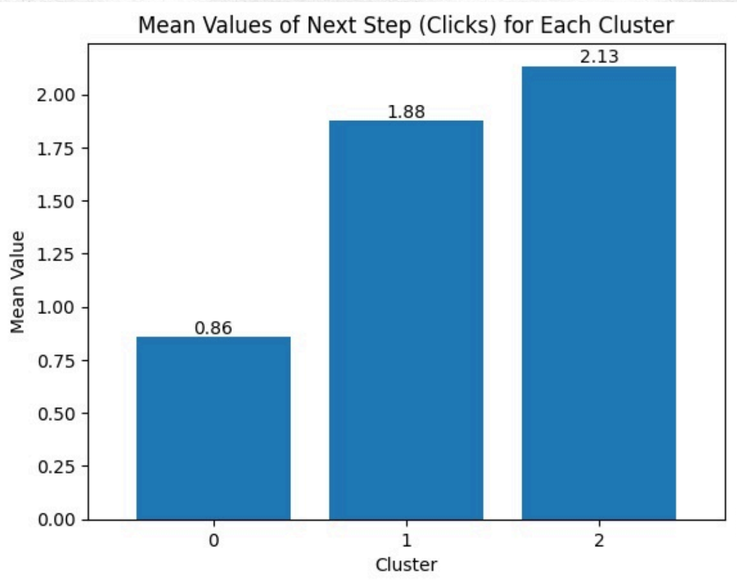
- More balanced engagement and fewer wasted clicks
- High success on Task 2 and moderate success on Task 3
- This group spent a considerable amount of time on FAQs and less on understanding the actual features and offerings
- Preferred structured outputs (Reports) but less reliant on technical tools, hinting towards a non-technical preference.

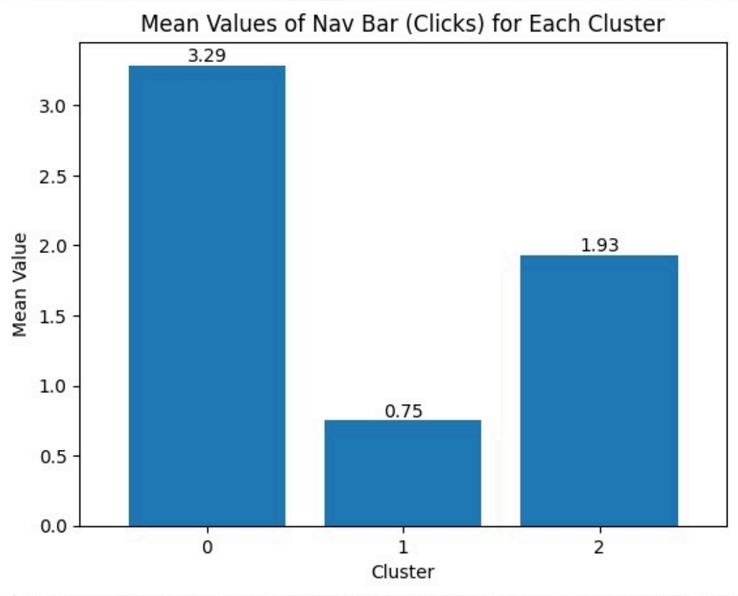
Cluster 2:

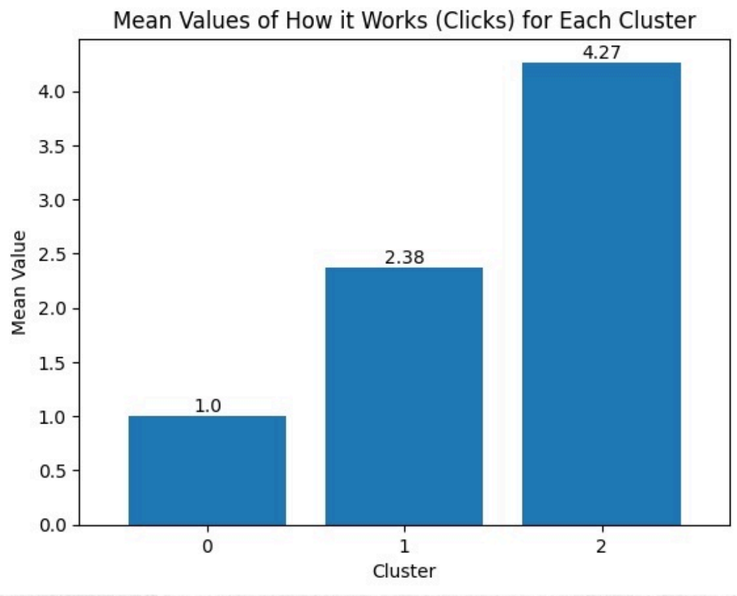
- Increasing cognitive load mid-task, then drop-off later
- Everyone succeeded at Task 2, but no one passed Task 3
- Spent majority of their time in How it Works & Features section trying to understand the offerings thoroughly
- Ended leaning toward guided or dedicated solutions after facing complexity

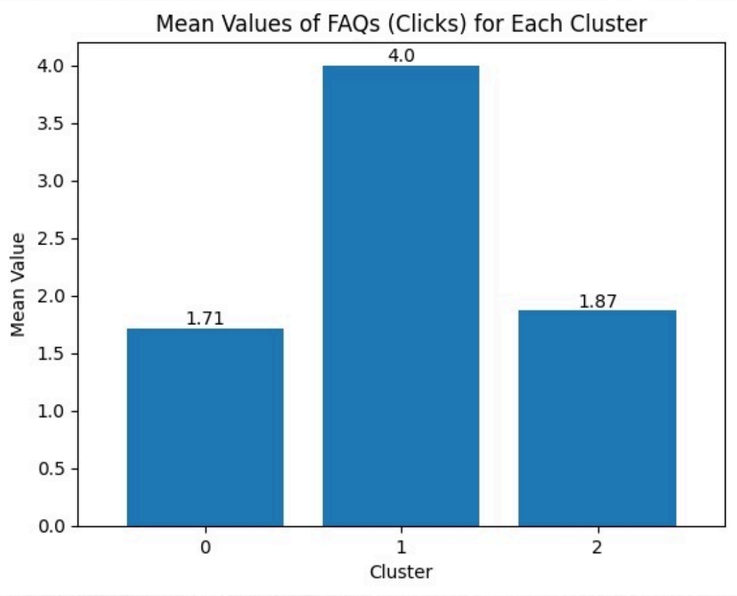

Mean Values of Total Clicks for Each Cluster
Mean Value
Cluster
11.29
9.12
15.27

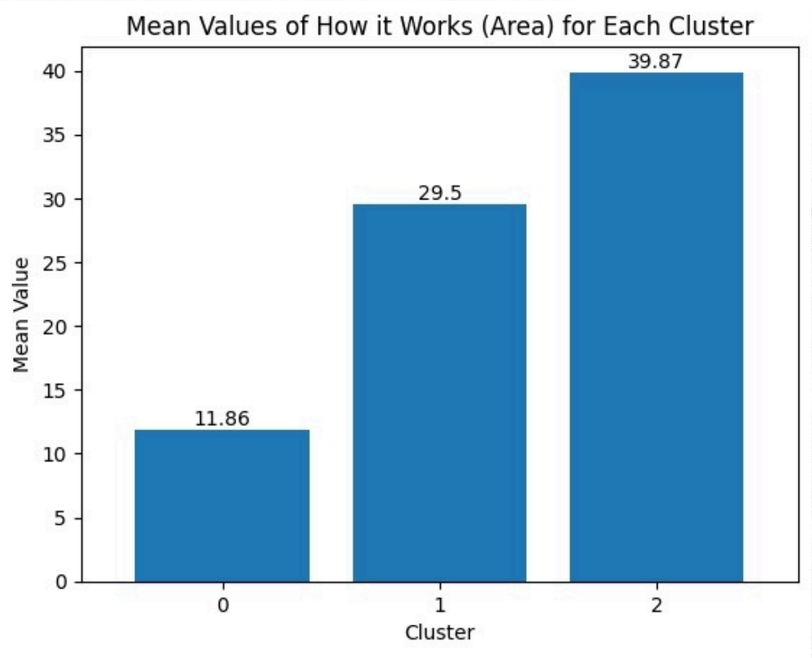

Mean Values of Dead Clicks for Each Cluster
Mean Value
Cluster
1.57
1.62
4.27


Mean Values of Next Step (Clicks) for Each Cluster
Mean Value
Cluster
0.86
1.88
2.13


Mean Values of Nav Bar (Clicks) for Each Cluster
Mean Value
Cluster
3.29
0.75
1.93

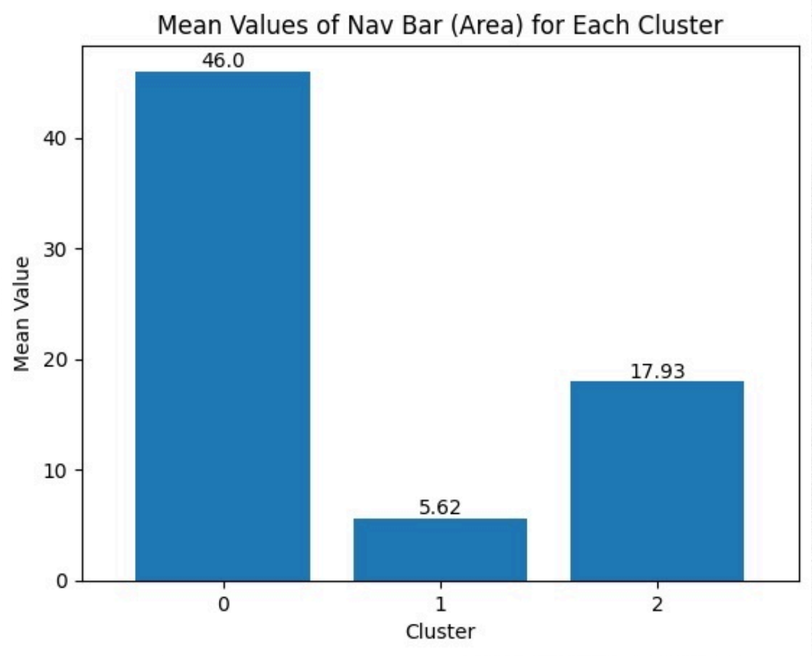

Mean Values of How it Works (Clicks) for Each Cluster
Mean Value
Cluster
1.0
2.38
4.27

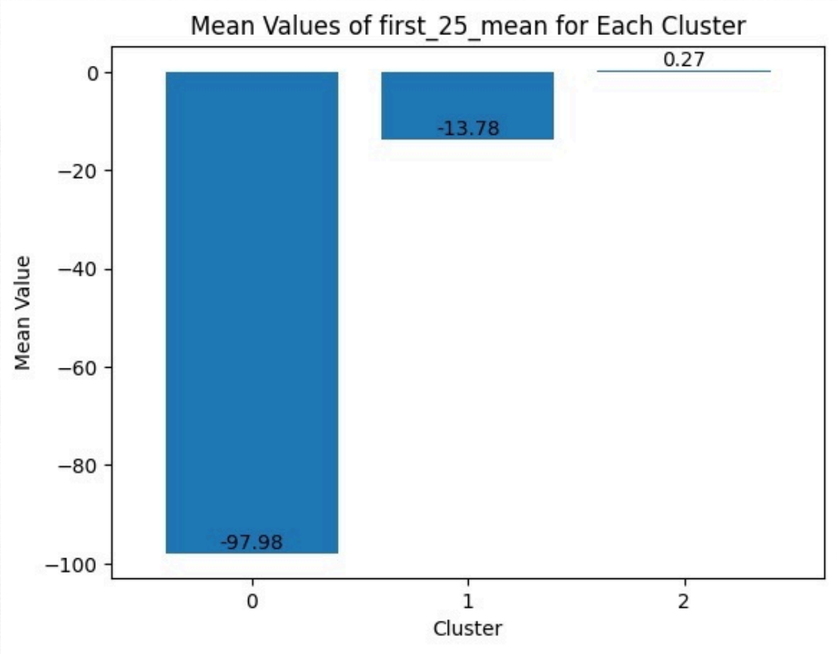

Mean Values of FAQs (Clicks) for Each Cluster
Mean Value
Cluster
1.71
4.0
1.87

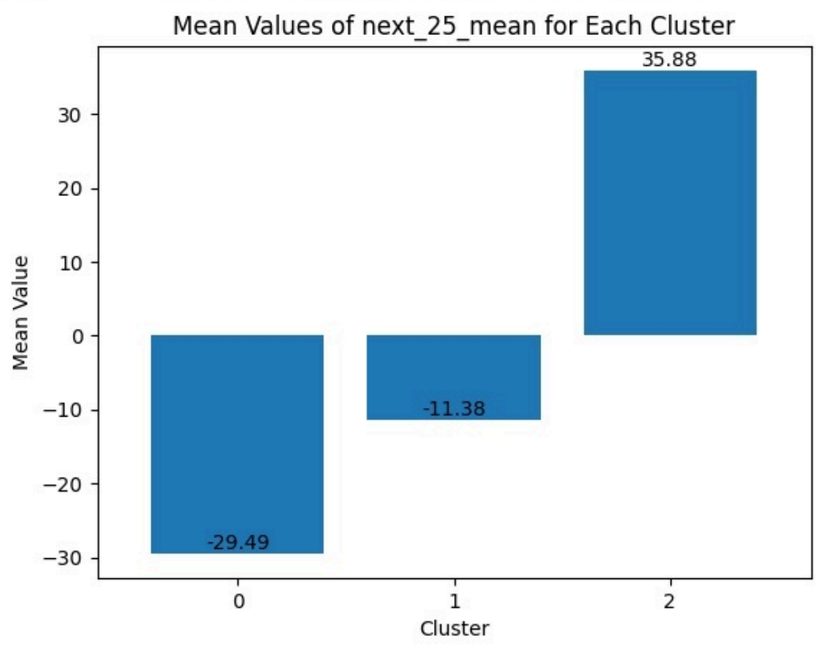

Mean Values of How it Works (Area) for Each Cluster
Mean Value
Cluster
11.86
29.5
39.87
0
1
2
0
5
10
15
20
25
30
35
40

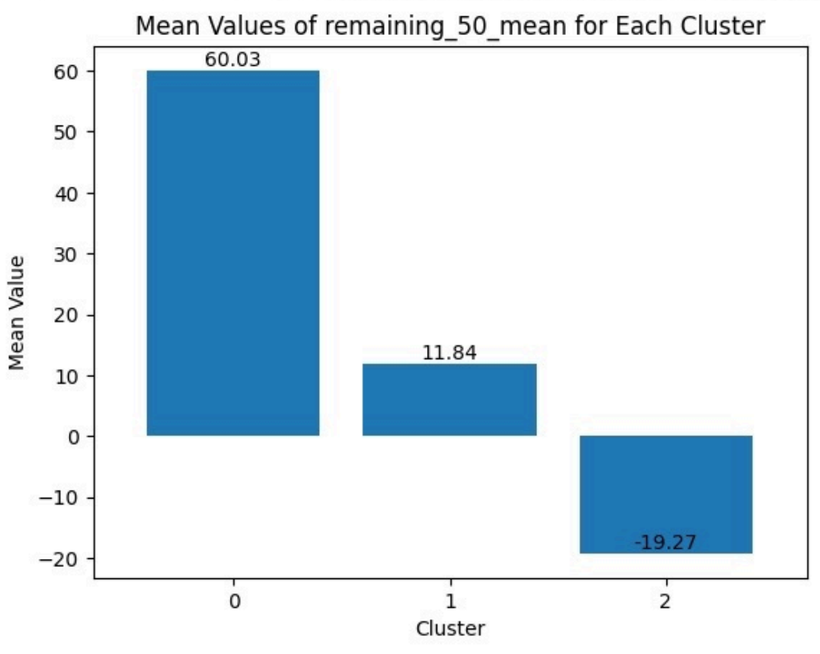

Mean Values of Features (Area) for Each Cluster
Mean Value
Cluster
20.71
4.62
24.73
0
1
2
0
5
10
15
20
25

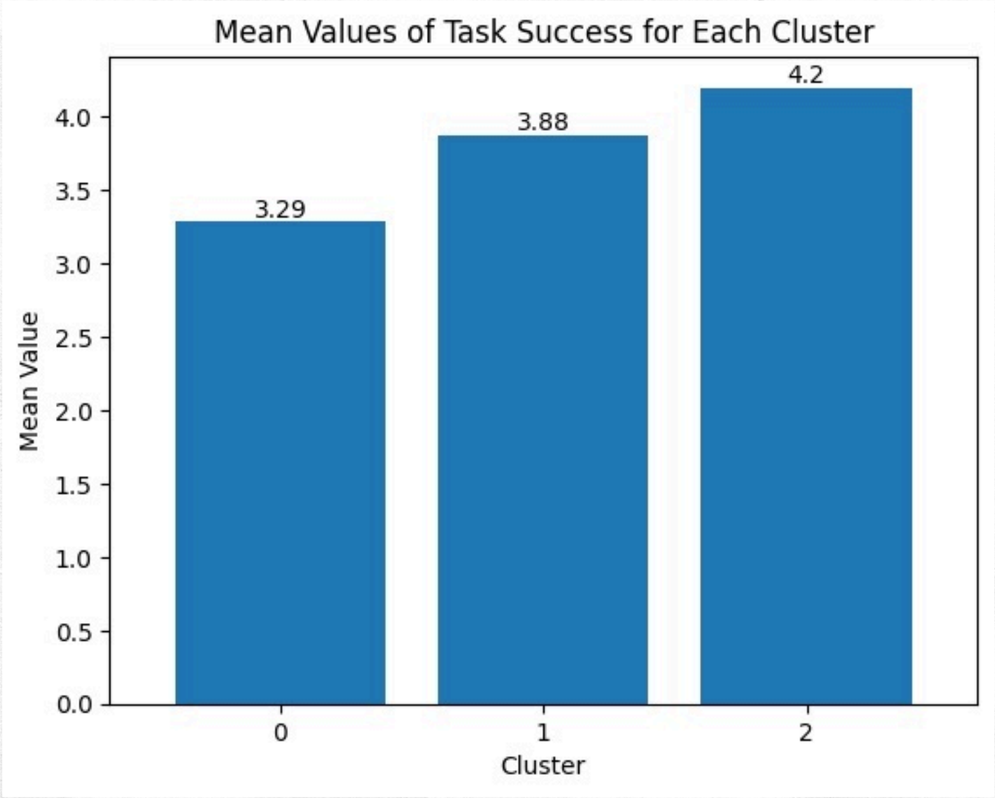

Mean Values of FAQs (Area) for Each Cluster
Mean Value
Cluster
21.43
41.88
17.07
0
1
2
0
5
10
15
20
25
30
35
40


Mean Values of Nav Bar (Area) for Each Cluster
Mean Value
Cluster
46.0
5.62
17.93
0
1
2
0
10
20
30
40


Mean Values of first_25_mean for Each Cluster
Mean Value
Cluster
0
1
2
0
−20
−40
−60
−80
−100
-97.98
-13.78
0.27


Mean Values of next_25_mean for Each Cluster
Mean Value
Cluster
0
1
2
30
20
10
0
−10
−20
−30
-29.49
-11.38
35.88


Mean Values of remaining_50_mean for Each Cluster
Mean Value
Cluster
0
1
2
60
50
40
30
20
10
0
−10
−20
60.03
11.84
-19.27


Mean Values of Task Success for Each Cluster
Mean Value
Cluster
3.29
3.88
4.2
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
0
1
2

Deployability

Real-World Deployment:

- The solution can be integrated into existing design and development pipelines.
- By continuously collecting user interaction data (clicks, navigation paths) and applying cognitive load estimations, designers can make iterative tweaks.
- Over time, this leads to a website or platform that dynamically adapts to user needs, refining its clarity, usability, and overall user satisfaction.

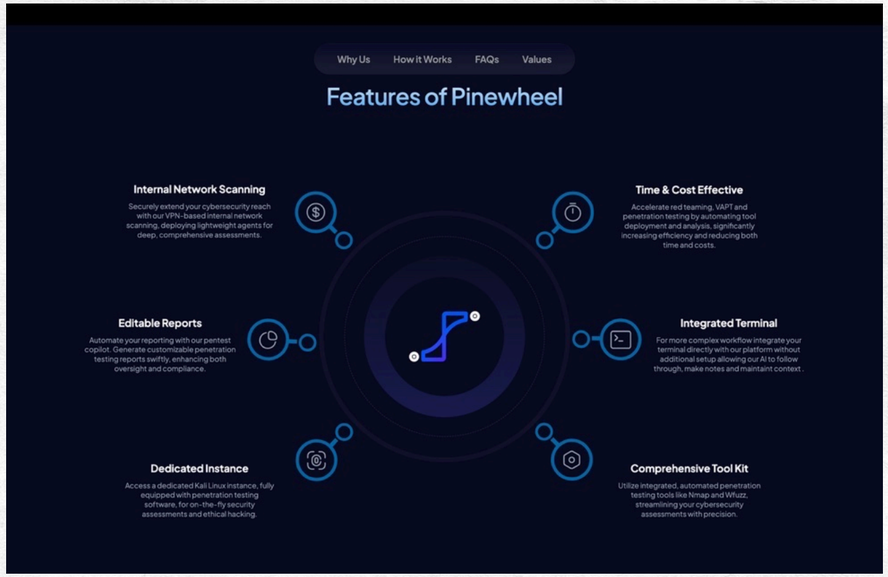
Challenges at Scale:

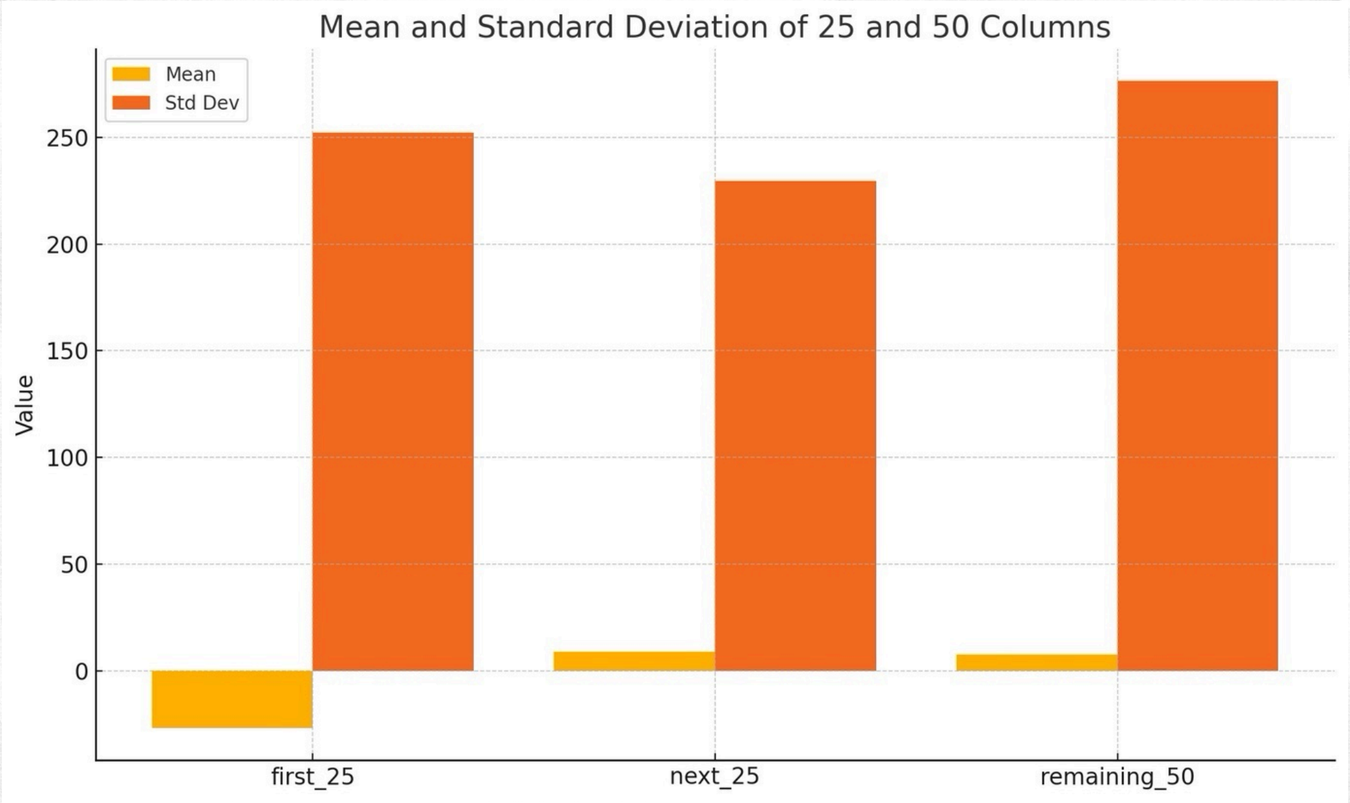
- **Data Privacy & Ethics:** Handling large volumes of user data requires strict adherence to privacy regulations and ethical standards.
- **Computational Resources:** Real-time analysis of cognitive indicators and user interactions may demand robust infrastructures (e.g., cloud computing).
- **Generalization Across Diverse Audiences:** Different user demographics, industries, and use cases may require context-specific adaptations, making it challenging to create a one-size-fits-all solution.
- **Complexity of Continuous Improvement:** Maintaining an ongoing feedback loop and prioritizing which design elements to optimize can become intricate as the user base and feature set grow.

IMPACT

IMPACT

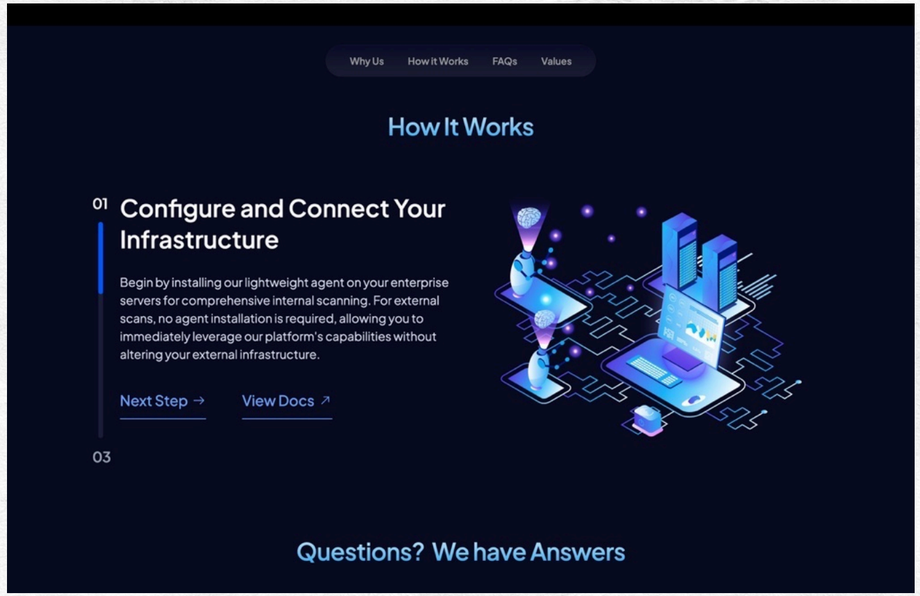
Higher cognitive load in the last task shows unclear / hard to understand information in the features section.

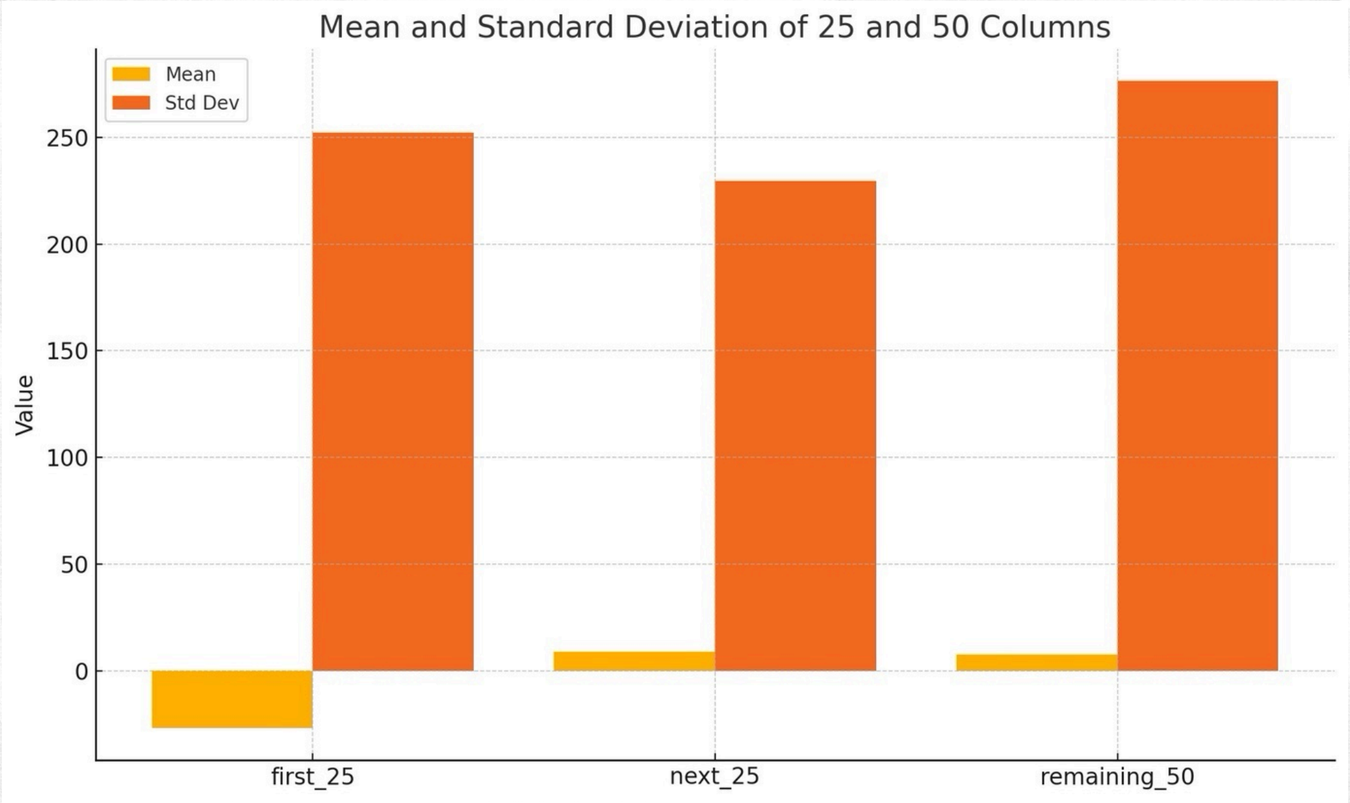




IMPACT

Lower cognitive load in the "How It Works" section suggests better user assimilation.





- **Personalized Onboarding:** Insights into user confusion patterns enable the creation of more tailored introductory content or onboarding tours for new users, ensuring individuals who initially struggle receive immediate guidance, reducing their initial cognitive load and frustration.

- **Adaptive Navigation Aids:** Recognizing clusters that stumble mid-way or later suggests introducing adaptive navigation aids, like context-sensitive tooltips or dynamic help prompts. These aids can appear when the system detects hesitancy or difficulty, helping users regain confidence and continue smoothly.

- **Customized Feature Placement:** For groups that show improved comfort over time, features can be phased in progressively. Initially, they might see simpler controls or curated feature sets, while more advanced or technical tools appear once they've demonstrated familiarity, reducing early overload and maintaining long-term engagement.

- **Better User Experiences Across Domains:** The framework can be applied to various industries—from e-commerce and e-learning platforms to government portals—making online interactions smoother and more inclusive.

Thank You!