Submitted Version of the manuscript; This preprint has not undergone peer review or any post-submission improvements or corrections. The Version of Record of this contribution is to appear in the Proceedings of the HCII'25

Case Study: Using EEG to Assess Cognitive Load in Infinite Interfaces

ANASTASIIA SATARENKO, MacPaw, Ukraine

In the realm of Human-Computer Interaction, cognitive load plays a crucial role in determining user satisfaction, product acquisition, and user retention. We are introducing a new type of interface - Infinite Interfaces, that dynamically adapts to user inputs and, therefore, offers personalized and efficient user experiences. However, the novelty of such interfaces may lead to increased cognitive effort, as most users are accustomed to traditional predetermined interfaces. This study seeks to evaluate the cognitive load associated with Infinite Interfaces using two approaches: subjective user ratings, which are simple and accessible, and EEG technology, providing an objective measure. The research aims to compare cognitive load in Infinite Interfaces versus traditional interfaces, using both methods to identify any similarities or differences in the results. Results indicate that while novel interfaces like Infinite Interfaces initially induce higher cognitive load, they offer long-term potential for improving user efficiency in complex tasks.

CCS Concepts: • **Human-centered computing** → Natural language interfaces; **Walkthrough evaluations**; *Usability testing*; *Scenario-based design*.

Additional Key Words and Phrases: EEG, Usability Testing, Natural Language Interfaces, Infinite Interface

1 INTRODUCTION

One of the central challenges in Human-Computer Interaction (HCI) is managing the cognitive load imposed on users by various new or advanced features in apps, which directly impacts their ability to navigate within the interface and complete tasks effectively. Cognitive load refers to the mental effort required to process information and achieve goals, drawing from limited working memory resources [6]. Managing this load is crucial, as excessive cognitive effort can lead to frustration and inefficiency.

According to a Nielsen Norman Group study, when talking about the user interface, cognitive load can be caused by 3 main factors: too many choices, too much thought required, and lack of clarity.

At the root of the Infinite Interface approach, we have the hypothesis that by taking out the unneeded and unused features from the interface, letting the user directly manipulate the interface by their input, and only proposing the results fit the user's needs - we will lessen the cognitive load the user experiences.

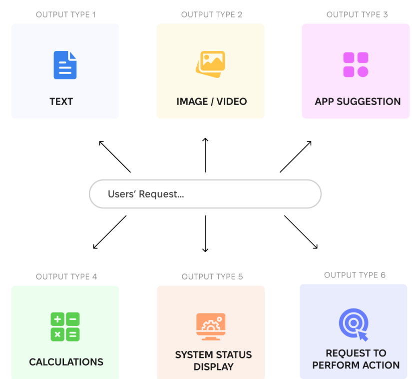
Infinite Interfaces are designed to tackle these challenges by minimizing visual clutter, removing irrelevant features, and providing users with full control over expressing their intent. The system then proposes relevant options based on user input, reducing the burden of decision-making and delivering actionable outcomes. This study investigates how the Infinite Interface design affects cognitive load, with a particular focus on measuring it through EEG technology and subjective user evaluations.

2 BACKGROUND

At the beginning of this year, when AI technology evolved enough to be easily incorporated into apps and websites, we started working on an AI assistant app that would help users with their everyday tasks and computer maintenance. The idea behind these interfaces was to use the users' inputs to create an interface itself, meaning no graphic interface will be shown unless the user writes their request. After the request is written, the interface generates a tailored solution for the user, whether it's an image, text, link to a website, app suggestion, or even the system's completion of the task. With this approach, the number of outputs is almost limitless, hence the name—infinite Interfaces. In this paper, we will

Author's address: Anastasiia Satarenko, MacPaw, Kyiv, Ukraine, an.sat@macpaw.com.

refer to Infinite Interfaces as "Oasis" as the title of this prototype. The interface can be represented schematically as shown in Figure 1a:



(a) Infinite Interface (Oasis) Schematic Visualisation

(b) CleanMyMac App with a Direct Manipulation interface

Fig. 1. Comparison of Oasis and CMM Interface Schematics

In this case study, we wanted to measure the cognitive load in an Infinite Interface, but we also needed to have "baseline" of cognitive load levels for the same tasks in a predetermined, dynamic graphic interface. For this case we chose a product "CleanMyMac", because it has the same features and the Infinite Interface product (deleting apps, malware cleaning, system management).

3 RELATED WORK

Cognitive Load Theory, introduced by John Sweller in 1988 [6] and building on earlier research by Miller [2], links cognitive load to task difficulty. Therefore, the effectiveness of a digital interface can be assessed by measuring the cognitive load it imposes. Common methods for evaluating cognitive load and interface quality include eye tracking, EEG, and usability testing with subjective user scores.

In cognitive psychology, cognitive load is the amount of working memory resources that are used to comprehend new information and keep track of one's goal. Cognitive Load Theory (CLT) [7] suggests that different pieces of information can be grouped into single units, called schemas, which can be automated. The main goal of a CLT-based interface is to create and automate these schemas. However, before information can be stored in long-term memory, it needs to be processed in working memory.

Short-term memory limitations dictate a whole range of other design guidelines [1]:

- quick response time so that the user doesn't have to keep in mind what task they are performing
- providing clarity within the interface to show what task is being performed
- the output of the interface should match semantically to the input

[1] https://www.nngroup.com/articles/minimize-cognitive-load/

There are numerous methods available for measuring cognitive load, each offering different levels of precision and applicability depending on the context. Among these, subjective rating scales [4] have emerged as the most commonly used approach in Cognitive Load Theory (CLT) research. These scales rely on users' self-reported perceptions of mental effort, providing a quick and practical way to gather insights into cognitive load during task performance.

Another common method for evaluating HCI systems has been Think Aloud Usability Testing[5]. In this approach, users perform tasks on applications while verbalizing their thoughts, and both verbal and behavioral data are collected to assess ease of use and efficiency. Usability testing also attempts to measure cognitive load, though indirectly, as slower TA behavior is often linked to higher cognitive load. However, since the TA method itself adds cognitive load, it may not be suitable for measuring cognitive load during demanding tasks. This paper suggests the need for alternative methods to assess cognitive load more accurately. It proposes using EEG, eye-tracking, and subjective user ratings as a more effective way to measure cognitive load in HCI systems.

A widely used tool in this area is the NASA-Task Load Index (NASA-TLX) [1], which has been adapted in many studies to assess cognitive load across various domains. The NASA-TLX evaluates mental, physical, and temporal demands, as well as effort, performance, and frustration, offering a comprehensive view of the user's perceived workload. Similarly, Paas' 9-point symmetrical mental effort rating scale [3] is another frequently employed method that directly measures the amount of mental effort a user feels they exerted on a given task.

These subjective rating scales are valued for their ease of use and reliability, providing researchers with consistent data on the cognitive demands of different tasks. While they depend on user perception, making them less objective than physiological measures, they offer a practical and validated approach for capturing cognitive load in real-time, although, for more contextual and fuller data, they are often complimented with objective assessments such as EEG and eye-tracking.

The most objective tool, EEG paired with eye-tracking, is a valuable instrument for assessing cognitive load in a non-intrusive and precise manner. It allows for the evaluation of brain waves at each step of a task, providing detailed insights into cognitive processes during task completion. By objectively measuring brain activity during users' interactions with the interface, EEG offers more reliable data compared to subjective assessments alone. When combined with eye-tracking, this method also enables a comprehensive understanding of the user's context, giving researchers a deeper and more accurate view of how users engage with the interface.

4 EXPLORATORY USABILITY TESTING

For our first study used of the CLT-based methods of Cognitive Load evaluation - the user's subjective score, so that we gather users' feedback on the interfaces, and also have a subjective data that we can later compare with EEG results.

We started with conducting a moderated user-testing with 2 different interfaces that have the same functionality: Infinite interface and Dynamic interface. Both interfaces are 2 different products that are system utilities and have to do with computer maintenance. One of them is an Infinite Interface (Oasis), and another one is CleanMyMac (CMM). As one of the tested products was still in the development phase and under NDA, we recruited volunteers from our company to participate in the testing. The target group was 7 people, 4 males and 3 females (age group 20-35 y/o), unfamiliar with both products we tested with them.

4.1 Testing Setup

The product and prototype we chose were chosen to compare due to similar functionality (Task 1, Task 2, and Task 3 are possible to complete in both apps), Oasis—Infinite Interface, and CleanMyMac (CMM)—Dynamic Interface.

Users were given 3 tasks, listed in Table 1. To assess the cognitive load when completing the tasks, we asked users for their subjective ratings. Users rated the simplicity of the completed task from 0 to 10, where 0 is the most difficult interaction, and 10 is the most intuitive and easy.

Task	**Process**	**Success Criteria**
Task 1: Delete the unused apps on your computer.	Find a way to see what apps are unused and delete them.	User deletes his unused apps.
Task 2: Check if your computer has any malware.	Find a way to scan your computer for viruses and check if you have any.	User can certainly say whether there are any viruses on his computer.
Task 3: Your computer is overheating and you don't know why. Find out the reason and fix it.	User needed to understand how to formulate the request himself or understand the root of the problem, find the solution himself, and complete several steps to resolve it.	User finds what would solve the problem in his opinion and is satisfied with the result.

Table 1. Task Processes and Success Criteria

4.2 Results

The results have shown, that the general score (the summary of all 3 tasks) was higher for Infinite Interfaces by 12% (see Figure 2 and Tables 2-3), meaning users think those types of interfaces were 12% easier for them to use.

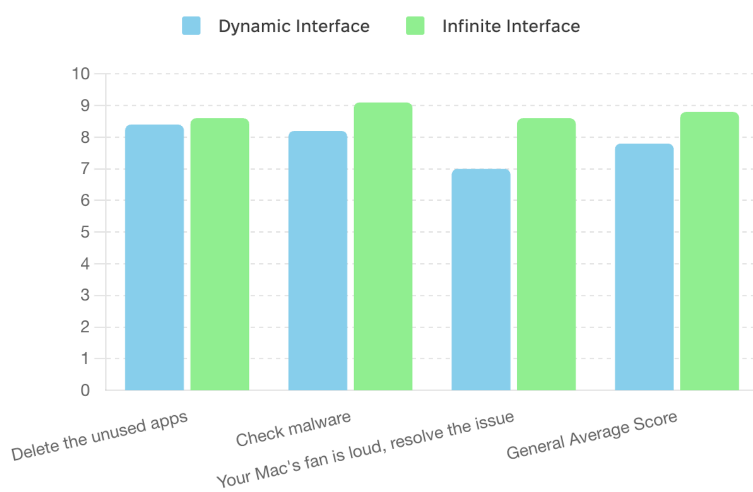


Fig. 2. Average Scores for Users' Subjective Simplicity Rating

What's more interesting, users rated Task 3 (the complex multi-step tasks) 23% better for Infinite Interfaces. We calculated the result based on the average score users gave to rate the simplicity of completing the given task (Task 3) in both interfaces.

	User 1	User 2	User 3	User 4	User 5	User 6	User 7	Average Score
Task 1	8	10	10	8	8	10	5	8.4
Task 2	7	8	8	9	10	9	7	8.2
Task 3	10	8	6	5	6	8	6	7.0
General Average Score								7.8

Table 2. User’s subjective satisfaction (CMM)

	User 1	User 2	User 3	User 4	User 5	User 6	User 7	Average Score
Task 1	9	10	8	7	9	8	9	8.6
Task 2	9	9	9	9	9	10	9	9.1 ↑
Task 3	7	10	8	10	9	6	10	8.6 ↑
General Average Score								8.8 ↑

Table 3. User’s subjective satisfaction (Oasis)

That gave us the idea that Infinite Interface (Oasis) brings more value for users in complex tasks, so we made the decision to assess cognitive load using a more objective and quantifiable approach and have the same testing conducted using EEG as a measuring tool, to see if the results will be confirmed. While the usability test provided valuable insights into users’ experiences and subjective perceptions, we wanted to complement those findings with a more data-driven assessment.

5 EEG AND EYE-TRACKING

By introducing a more objective method of measuring cognitive load—such as using EEG, eye-tracking, or physiological measurements, we aimed to determine whether the cognitive load users experienced aligned with their subjective score during the usability test.

This approach allows us to see if there is a correlation between users’ self-reported difficulty and their actual cognitive load during tasks. Additionally, it provides us with a more accurate understanding of how demanding certain tasks or interface elements are, enabling us to identify specific areas where the interface may need to be refined.

EEG captures brain activity across multiple indicators, such as attention, cognitive processing, and emotional response during each task. Eye-tracking was also employed to assess engagement and fixations during the users’ interaction with the apps. After the test, the users also took part in a user interview and a survey.

5.1 Testing Setup

We recruited 20 participants (10 men and 10 women, aged 20-30) for the study. The users were selected by a 3-rd party and compensated for taking part in the test. The test was conducted offline, in the 3rd party office, with the EEG equipment and a computer with the eye-tracking camera installed (see Figure 3). All of the users were unfamiliar with both interfaces and were using the tested apps for the first time.

They completed the same tasks using the same two interfaces: an Infinite Interface and a Dynamic interface. Tasks included app deletion, virus removal, and system diagnostics on a Mac (Task 1, Task 2, Task 3), that were listed before. For the test, we have an Infinite Interface, our prototype with an internal name "Oasis" (further on, "Oasis" will be referring to Infinite Interface), and a Dynamic Interface, which is the CleanMyMac app, that further on will be named "CMM".

Fig. 3. Conducting the usability test using the EEG

5.2 Results

During completion of all the tasks, we measured the time taken to complete the task and also the brain activity, including emotional background, engagement, attention level, and cognitive load, but in this case study, we will focus solely on cognitive load and time measurement, as only those two contribute to the topic.

In the time-related tasks, "SD" stands for Standard Deviation.

Task 1: Delete the unused apps on your computer.

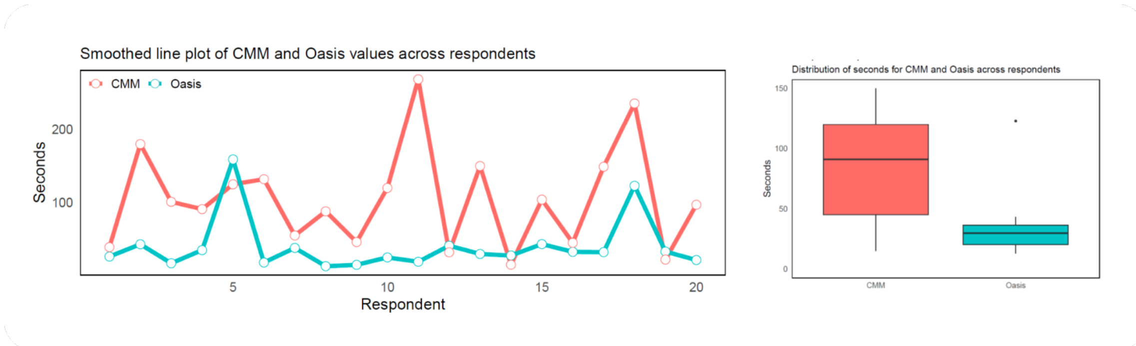
The average duration of the task: Oasis: 39 sec. (SD=37.3), CMM: 109 sec. (SD=69.1), see Figure 4a.

Thinking about an action when using the CMM is much less demanding on the brain, because the respondent has a number of options to choose from. The brain load when uninstalling programs is similar, which may indicate the equivalence of the final process but greater complexity of the path to it (in Oasis), see Figure 6.

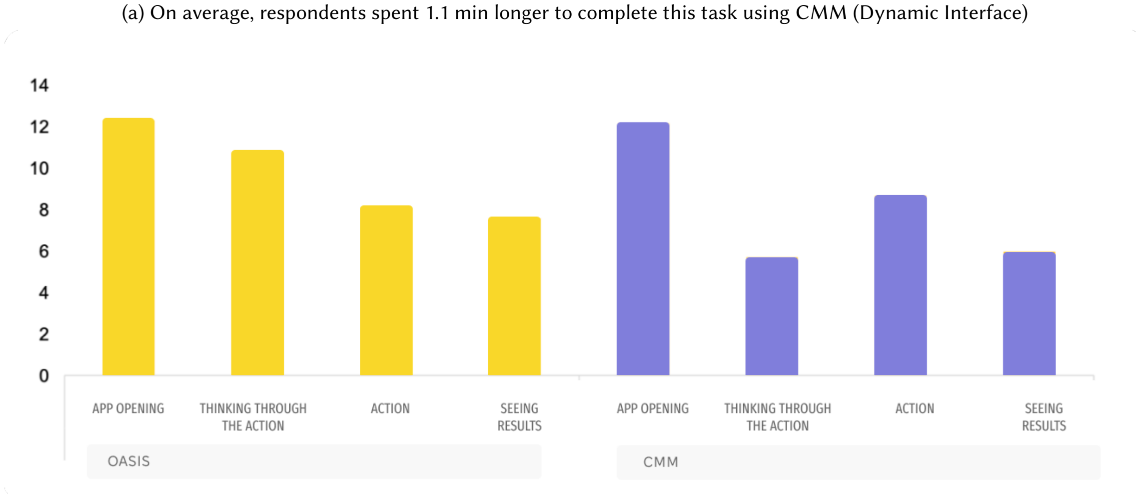
Task 2: Check if your computer has any malware.

The average duration of the task: Oasis: 70 sec. (SD=77), CMM: 21 sec. (SD=15), see Figure 5a.

When working with CMM (see Figure 5b), there is a noticeable decrease in cognitive load as users progress through each step. The interaction becomes progressively easier because the familiar interface provides a clear, built-in sequence of actions that guide the user effectively. This structured approach contrasts with Oasis, an Infinite Interface, where users experience a higher cognitive load, especially when it comes to interpreting the output or 'seeing the results.' In Oasis, users not only need to understand the outcome but also evaluate whether it aligns with their original request,



(a) On average, respondents spent 1.1 min longer to complete this task using CMM (Dynamic Interface)



(b) Cognitive Load Measurement: Task 1

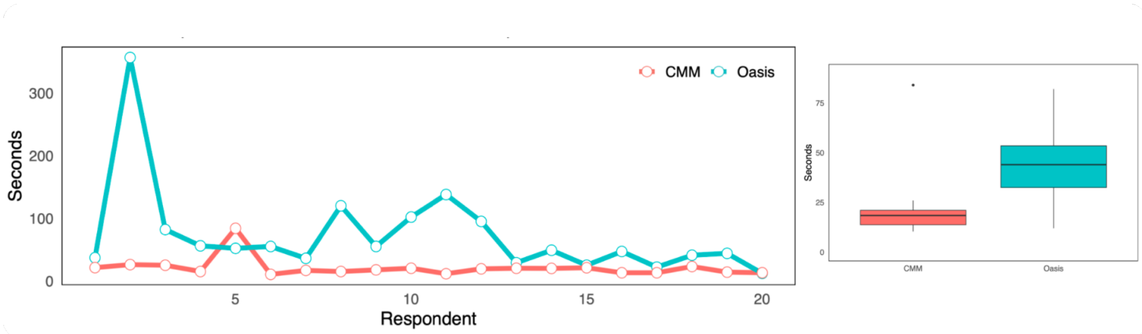
Fig. 4. Stacked wide plots in ACM single-column format

which can lead to confusion. This confusion often stems from differences in terminology, such as when a user inputs a search term like ’virus,’ but the result is labeled as ’malware.’ This subtle shift in semantics can cause users to question whether the result is truly relevant to their initial query, adding to the mental effort required.

Task 3: Your computer is overheating and you don’t know why. Find out the reason and fix it.

The average duration of the task: Oasis: 99 sec. (SD=148), CMM: 202 sec.(SD=191), see Figure 6a.

Since this behavior is not intuitive for users, it requires a significant amount of mental resources for them to process both their initial request and the actions they need to take. Users must put extra cognitive effort into understanding how the system works and determine whether the output meets their expectations. However, this trend can be influenced and potentially improved through repeated experience with the program. As users become more familiar with the interface and its logic, their interactions can become smoother, and they may require less brain load to navigate the system or evaluate results. Over time, this familiarity can help reduce the cognitive load, making the process more intuitive.



(a) Users were 49 sec. faster to complete this task in CMM

(b) Cognitive Load Measurement, Task 2

Fig. 5. Stacked wide plots in ACM single-column format

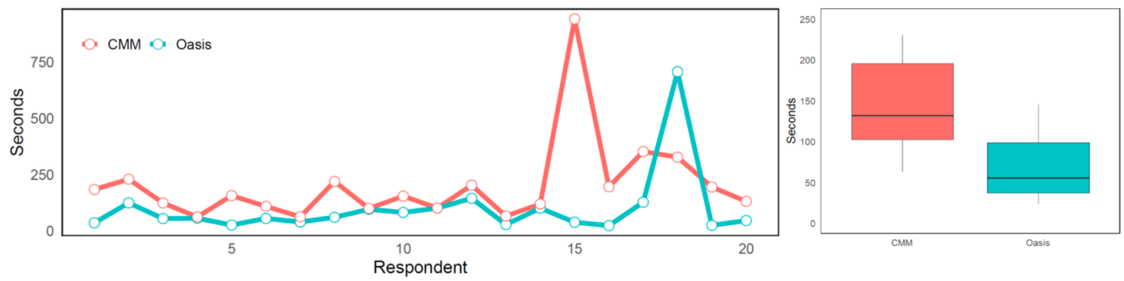
The "seeing results" column in the Figure 6b is not shown and evaluated because the success criteria for this task weren't the same for every user and was rather subjective: the task was considered finished once the user was happy with the result. Therefore, for this case, we focused on measuring the process of getting to the desired outcome.

5.3 Exit survey

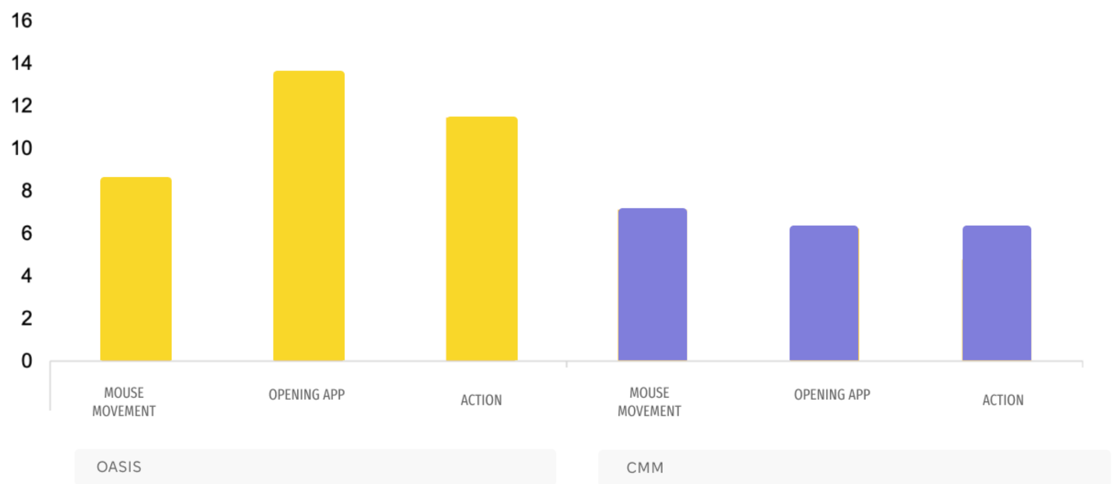
Apart from measuring cognitive load during users' task completion, we also conducted an exit survey, where we measured users' cognitive load when asked about tasks they're willing to delegate to their utility app / AI assistant. As we tested a certain type of interface with the users, we also wanted to validate, what features bring the most cognitive load on users, so that we pay extra attention to them and

Numbered task list:

(1) Clean up unnecessary files
(2) Access to calendars, emails, Zoom, Google Meet



(a) On average, respondents spent 1.2 min longer to complete this task using CMM



(b) Cognitive Load Measurement, Task 3

Fig. 6. Stacked wide plots in ACM single column format

(3) Protection against viruses
(4) Email management and replying to emails
(5) Downloading videos
(6) Manage tasks and projects
(7) Conversion, editing, document management
(8) Password manager/generator

In the figure below, each number corresponds to the numbered task list above. The highest rates of brain load were recorded when thinking about giving an AI assistant the ability to manage email and passwords, which could possibly mean that users are not yet comfortable with delegating personal matters to the assistant. These tasks involve high levels of trust and control over personal data, which can evoke anxiety or hesitation in users, leading to increased

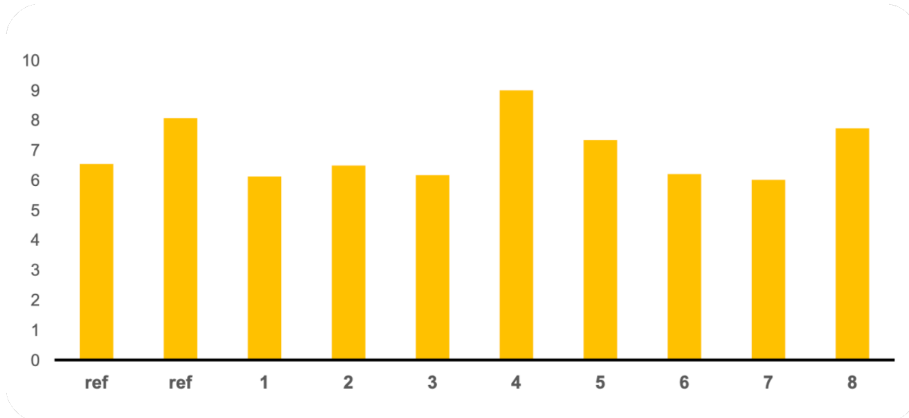


Fig. 7. Exit Survey Arithmetic Mean for Cognitive Load when delegating tasks to the assistant / AI

mental effort when considering such delegation. However, by knowing this, we can better understand which aspects of the interface design and product features contribute to user stress and where improvements can be made.

6 DISCUSSION

The initial cognitive load that users experience when working with Oasis can largely be attributed to its unfamiliar and non-traditional interface, which demands that users put more mental effort into understanding and navigating it. This higher cognitive load is typical when users first interact with complex systems that do not align with their existing mental models or prior experiences. However, as users gradually become more familiar with the Infinite Interface, the time required to complete even the most complex tasks decreases. This suggests that, despite the learning curve, Infinite Interfaces offer potential long-term benefits in terms of efficiency and ease of use. These observations are supported by EEG data, which show a clear reduction in cognitive load as users acclimate to the interface. As they gain more experience, their brain activity reflects less strain and greater comfort in interacting with the system, reinforcing the idea that while the interface may initially be challenging, repeated use leads to more streamlined and cognitively efficient interactions.

After measuring cognitive load using EEG during the interaction with different interfaces, we can draw several key conclusions:

(1) **Subjective user scores align with EEG data but are not absolute indicators.** Our findings support the use of subjective user scores as a valid measurement for interface usability; however, EEG data offers a more nuanced picture. While participants reported certain experiences as more intuitive or simple, the cognitive load measured through EEG revealed additional complexities. This indicates that while self-reported ease-of-use is important, it can sometimes mask the cognitive effort required, suggesting that relying solely on subjective metrics may not capture the full scope of user experience.

(2) **Users often confuse speed with cognitive simplicity**: One of the most interesting insights from this study is that users tend to associate speed of task completion with simplicity, even when their cognitive load is higher. For instance, when asked which interface was simpler to use, participants rated the Oasis (an Infinite Interface) 12% higher for general tasks and 23% higher for more complex tasks (e.g., Task 3) than a Dynamic Interface.

Despite the fact that EEG readings indicated significantly higher cognitive load during the use of the Infinite Interface, participants still perceived it as simpler. This disconnect could be due to the speed with which they achieved their results. In this case, rapid task completion seemed to compensate for the higher mental effort, leading users to equate speed with ease of use. Therefore, this points to a potential cognitive bias where users perceive fast interfaces as less mentally taxing, even when the opposite is true.

(3) **There are request formulation difficulties in Infinite Interfaces.** During the test, we discovered the most popular issue that users had while interacting with the Infinite interface was about formulating the request ("What do I type here?", "How do I formulate a question?"), which is presumably caused by a lack of mental models when interacting with such interfaces because there are very few interfaces with similar functionality or a lack of clarity on the capabilities of the program and clear instructions about the request forming.

(4) **Delegating personal tasks to an assistant or utility increases cognitive load**: Another significant observation is that users experienced heightened cognitive load when they had to delegate personal tasks to an assistant or utility interface. This increase in brain activity suggests that even though the task is being automated, users may still feel an implicit responsibility or anxiety when handing over control to a system. The mental strain associated with trusting technology to perform personal or critical tasks could be a contributing factor. It also implies that while automation is designed to ease user burden, transitioning from manual to automated task completion might require more cognitive adjustment than expected.

Findings from this research directly affected the product development process and approaches we were taking when working on a new app with an Infinite Interface. We understood at what stages do we overload the users' brain and developed a plan to assess those pain points.

7 CONCLUSION

This case study highlights the effectiveness of EEG in assessing cognitive load in interface designs. While Infinite Interfaces like Oasis initially demand higher cognitive resources, they present opportunities for enhanced user efficiency in the long run. By tracking cognitive load in real-time, EEG allows researchers and designers to pinpoint exactly where users struggle or experience mental fatigue, enabling more targeted improvements in the design process. This method can help design user-friendly solutions that maintain the innovative potential of Infinite Interfaces. Monitoring cognitive load at various stages of product development is crucial, as users, particularly new ones, may abandon a product if it is too difficult to understand or use. This can result in substantial retention and conversion losses for companies, making it essential to address and minimize brain load early on in the design. By proactively reducing cognitive barriers, designers can ensure that users remain engaged with the product, improving both user satisfaction and long-term success. Future research should explore methods for reducing initial cognitive load in Infinite Interfaces while preserving their innovative advantages.

8 ACKNOWLEDGMENTS

We thank the Armed Forces of Ukraine for providing security to complete this work. We thank Nataliia Stulova for providing assistance in preparing this manuscript.

REFERENCES

[1] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*, Peter A. Hancock and Najmedin Meshkati (Eds.). Human Mental Workload, Vol. 52. North-Holland, 139–183. https:

//doi.org/10.1016/S0166-4115(08)62386-9

[2] George A. Miller. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63, 2 (1956), 81–97. https://doi.org/10.1037/h0043158

[3] Fred Paas, Juhani E. Tuovinen, Huib Tabbers, and Pascal W. M. Van Gerven. 2003. Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. In *Cognitive Load Theory*. Routledge.

[4] Annett Schmeck, Maria Opfermann, Tamara Van Gog, Fred Paas, and Detlev Leutner. 2015. Measuring cognitive load with subjective rating scales during problem solving: differences between immediate and delayed ratings. *Instructional Science* 43, 1 (Jan. 2015), 93–114. https://doi.org/10.1007/s11251-014-9328-3

[5] Michael J. Smith, Richard John Koubek, Gavriel Salvendy, and Don Harris. 2001. *Usability Evaluation and Interface Design: Cognitive Engineering, Intelligent Agents, and Virtual Reality*. CRC Press.

[6] John Sweller. 1988. Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science* 12, 2 (April 1988), 257–285. https://doi.org/10.1207/s15516709cog1202_4

[7] John Sweller. 2011. CHAPTER TWO - Cognitive Load Theory. In *Psychology of Learning and Motivation*, Jose P. Mestre and Brian H. Ross (Eds.). Vol. 55. Academic Press, 37–76. https://doi.org/10.1016/B978-0-12-387691-1.00002-8