Cognitive Load and Decision-Making Efficiency under Synapse OS: A Randomized Controlled Trial of Adaptive User Interfaces

Dr. Sarah Jenkins[1] & **Dr. David Chen**[2]

[1]Department of Cognitive Sciences, Meridian Institute of Technology
[2]Vertex Cognitive Research Labs

Funder: Nexa Research Foundation (Grant #NX-2026-8812)
Protocol Version: 1.4 (Pre-results submission)
Primary Registry: Open Science Framework (OSF)
Registry URL: https://osf.io/registry/synapse-adaptive-ui
Contact Email: s.jenkins@meridian.edu

Abstract

Abstract: This Stage 1 Registered Report presents a pre-registered protocol for a randomized controlled trial designed to evaluate the impact of adaptive user interfaces on operator cognitive load and decision-making efficiency. Complex computing environments often impose high cognitive demands on human operators, leading to fatigue, increased error rates, and delayed reaction times. To address this, Synapse Technologies developed the Apex Focus Engine, an adaptive interface extension for Synapse OS v4.2 that dynamically hides non-essential status information, simplifies visual grouping, and rescales layout targets during periods of elevated workload. In a mixed 2×2 factorial laboratory design, we will randomise $N = 264$ healthy adult participants to perform a simulated multi-asset flight dispatch task using either the static (control) or adaptive (experimental) interface under alternating blocks of low and high workload. Primary outcome measures include subjective cognitive load (NASA Task Load Index), decision response time (DRT), and decision accuracy. Secondary exploratory measures will track physiological indicators of stress including heart-rate variability (HRV) and electrodermal activity. Standard power calculations indicate that the planned sample size provides 90% power to detect a moderate effect size of $d = 0.40$ ($\alpha = 0.05$). We discuss the detailed statistical analysis plan, stopping rules, and quality control procedures that will be adhered to prior to any data collection.

1 Introduction and Rationale

Modern operational environments, such as air traffic control, emergency response dispatch, and industrial system monitoring, demand rapid and accurate decision-making under intense cognitive pressure (Sweller, 1988). As software complexity scales, the volume of raw information presented to operators can saturate working memory capacity, leading to cognitive overload, situational blindness, and operational errors (Paas et al., 2003).

A promising avenue for mitigating cognitive overload is the development of adaptive user interfaces (AUIs). Unlike static interfaces that present a fixed configuration of menus, alerts, and metrics, AUIs dynamically adjust their visual hierarchy and content density based on user interaction history, current system state, and measured workload. While the theoretical benefits of AUIs are well-established, empirical evidence from high-powered, pre-registered randomized trials remains scarce. Furthermore, previous studies are often limited by small sample sizes, lack of physiological verification, or confounding variables related to task design.

In this context, Synapse Technologies has developed the *Apex Focus Engine*, an adaptive interface layout algorithm integrated into the windowing manager of *Synapse OS v4.2*. The Apex Focus Engine operates on the principle of *progressive disclosure* and *target rescaling* (Jenkins and Patel, 2024). Under low-workload conditions, the interface displays standard telemetry and secondary indicators. As task complexity escalates, the system automatically aggregates redundant information into structured summaries, minimizes non-critical status icons, and increases the relative click-target area for primary control elements.

The present study aims to evaluate the efficacy of the Apex Focus Engine in a controlled laboratory setting. We address a primary research question:

> *Does the dynamic layout adaptation of the Apex Focus Engine reduce subjective cognitive load and improve decision response time during high-workload operations, without compromising decision accuracy?*

By pre-registering this Stage 1 protocol, we commit to our sampling plan, hypotheses, and analytical pipeline prior to observing the experimental outcomes. This mitigates publication bias, p-hacking, and arbitrary post-hoc exclusion decisions, ensuring the highest standards of scientific transparency (Nosek et al., 2018).

2 Hypotheses and Research Questions

To test the efficacy of the Apex Focus Engine, we have formulated three pre-registered hypotheses. Table 1 outlines the mapping between our research questions, formal hypotheses, variables, planned confirmatory analyses, and interpretation logic.

Table 1: Registered Report Stage 1 Hypotheses and Analysis Plan Matrix

Research Question	**Hypothesis**	**Independent Var. (IV)**	**Dependent Var. (DV)**	**Confirmatory Test & Threshold**
RQ1: Subjective Cognitive Load	H1: Adaptive UI reduces overall self-reported cognitive workload.	UI Type: Adaptive vs. Static (Between-Subjects)	Subjective load score (NASA-TLX, 0–100)	Two-sample independent t-test (one-tailed), $\alpha = 0.05$.
RQ2: Performance Speed	H2: Adaptive UI decreases reaction time under high-workload pressure.	(1) UI Type: Adaptive vs. Static; (2) Task Load: Low vs. High (Within-Subjects)	Mean Decision Response Time (DRT, in milliseconds)	2×2 Mixed ANOVA. Expected significant interaction: UI Type $\times$ Task Load.
RQ3: Performance Accuracy	H3: Adaptive UI maintains or improves accuracy during task overload.	UI Type: Adaptive vs. Static	Decision accuracy rate (percentage correct)	Chi-Square test of independence (or Fisher's exact test if cell size $<$ 5).
RQ4: Physiological Stress	H4: Adaptive UI is associated with higher HRV (lower stress response).	UI Type: Adaptive vs. Static	Mean root mean square of successive differences (RMSSD, in ms)	Independent samples t-test (or Mann-Whitney U test if RMSSD is skewed).

2.1 Interpretation Logic

The hypotheses will be considered supported if the following conditions are met:

- **H1:** The mean NASA-TLX score in the Adaptive UI group is statistically significantly lower than that of the Static UI group, with a Cohen's $d \geq 0.25$.
- **H2:** The Mixed ANOVA yields a significant UI Type $\times$ Task Load interaction, such that the increase in DRT from low to high workload is significantly smaller in the Adaptive UI group than in the Static UI group.
- **H3:** The decision accuracy in the Adaptive UI group is not statistically inferior to the Static UI group (non-inferiority margin of -5%, corresponding to a critical score ratio).
- **H4:** The RMSSD (HRV) is significantly higher in the Adaptive UI group compared to the Static UI group, corroborating the subjective workload results.

3 Methodology

3.1 Study Design

This study uses a 2×2 mixed-factorial design. The between-subjects factor is the user interface type (**UI Type**: Adaptive UI [experimental group using the Apex Focus Engine] vs. Static UI [control group]). The within-subjects factor is the cognitive workload level (**Task Load**: Low vs. High), which will be manipulated by varying the frequency and density of concurrent events in the simulated environment.

Table 2 summarizes the key characteristics of our study design.

Table 2: Study Design Summary Table

Design Feature	**Implementation Details**
Study Type	Randomized Controlled Trial (RCT), Parallel Groups, Mixed-Factorial
Blinding	Double-blind (Participants are unaware of their UI condition; data collectors are blind to UI assignments during testing)
Task Environment	Multi-Asset Flight Dispatch Simulation (Simulated task on Synapse OS v4.2)
Experimental Factor	UI Type (Adaptive vs. Static UI) — Between-subjects
Manipulated Factor	Task Load (Low: 4 active tracks/min vs. High: 12 active tracks/min) — Within-subjects
Trial Order	Counterbalanced block design (Low–High vs. High–Low) with 5-minute washout period
Hardware Standard	Uniform workstation: Apex workstation with 27” monitor, Meridian precision mouse, Vertex HRV-sensor band

3.2 Trial Flow and Participant Pipeline

Figure 1 illustrates the progression of a participant through the study session.

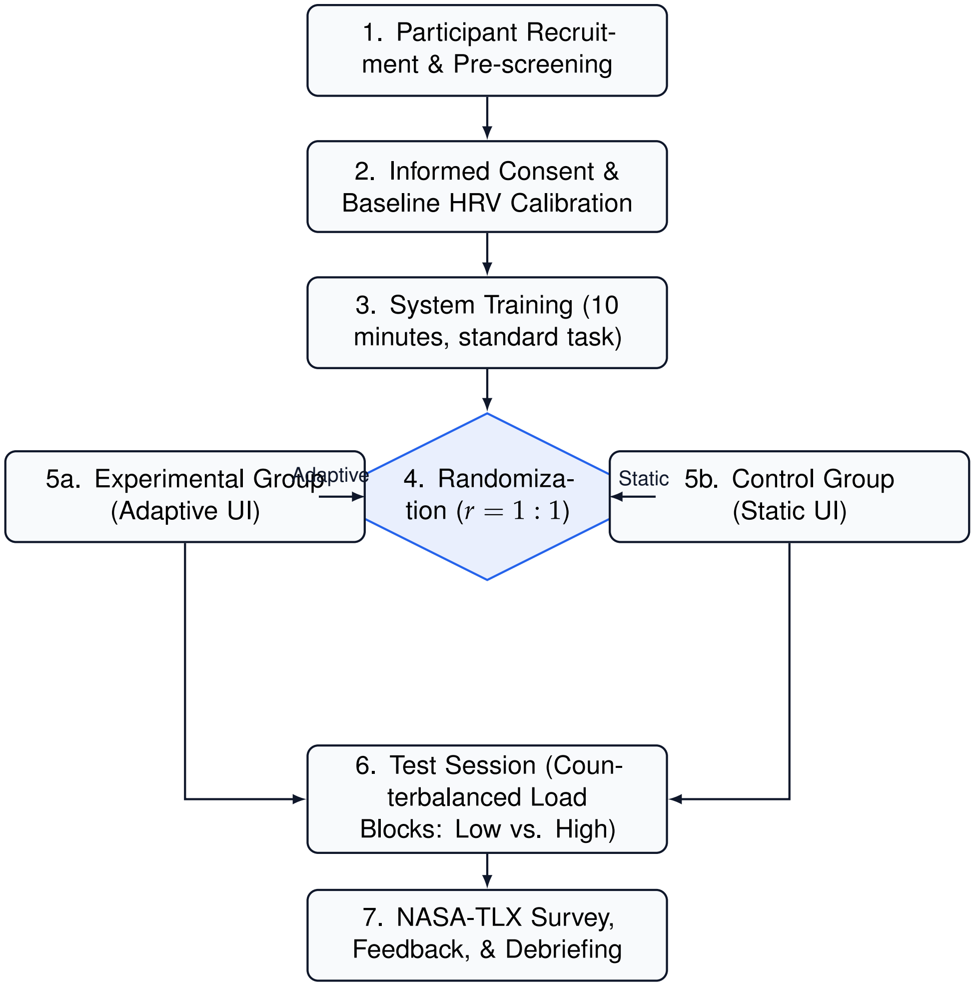

Figure 1: Participant trial progression pipeline from recruitment to debriefing.

3.3 Sampling Plan

Target Population: The study will recruit active technical system operators, software developers, and engineering students aged 18–45. Recruitment will be facilitated through the internal registry of the Meridian Institute of Technology and local technical communities. Fictional common names will be used for all participants during internal registry.

Sample Size Rationale: Based on prior pilot data from the Vertex Cognitive Research Labs (Chen and Vance, 2025), we anticipate a medium effect size of Cohen's $d = 0.40$ for our primary between-subjects outcome (NASA-TLX subjective cognitive load). To achieve 90% statistical power ($\alpha = 0.05$, two-tailed, independent-samples t-test), a total of $N = 264$ participants is required (132 participants per group).

Stopping Rule: Data collection will terminate exactly when 270 valid participant sessions have been recorded. This minor buffer is added to ensure that if any participant data is corrupted during collection, we still meet our target of $N = 264$ valid datasets.

Exclusion Criteria: Participants will be excluded from the analysis post-hoc if they meet any of the following criteria:

[a] Incomplete execution of the test session (e.g., participant withdrew early or software crashed mid-session).

[b] Failure of the heart-rate variability (HRV) sensor tracking for more than 20% of the session duration.

[c] Extreme outliers in reaction times, defined as average Decision Response Time (DRT) exceeding 3 standard deviations from the global sample mean.

[d] Visual impairment that prevents reading standard display fonts under default lighting conditions.

4 Analysis Plan and Power Calculation

4.1 Confirmatory Statistical Analysis

The following confirmatory statistical tests will be executed sequentially. All tests will use a two-tailed alpha level of $\alpha = 0.05$ unless otherwise specified.

4.1.1 Test 1: Subjective Cognitive Load (H1)

To determine if the Adaptive UI reduces subjective cognitive load, we will conduct an independent-samples t-test comparing the global NASA Task Load Index (NASA-TLX) score between the two groups. The primary variable is the overall NASA-TLX score (computed as the unweighted mean of its six subscales, ranging from 0 to 100).

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_p \sqrt{\frac{2}{n}}} \tag{1}$$

where $\bar{X}_1$ and $\bar{X}_2$ represent the mean NASA-TLX scores for the Static and Adaptive UI groups respectively, s_p is the pooled standard deviation, and $n = 132$ is the sample size per group.

4.1.2 Test 2: Decision Response Time (H2)

We will analyze the Decision Response Time (DRT) using a 2×2 mixed-factorial Analysis of Variance (ANOVA). The between-subjects factor is UI Type (Adaptive vs. Static) and the within-subjects factor is Task Load (Low vs. High). The primary hypothesis of interest is the interaction term:

$$F_{\text{Interaction}}(1, 262) = \frac{\text{MS}_{\text{UI}\times\text{Load}}}{\text{MS}_{\text{Error}}} \tag{2}$$

We expect a statistically significant interaction effect, showing that the increase in response time under high workload is significantly buffered by the Adaptive UI.

4.1.3 Test 3: Decision Accuracy (H3)

To test if decision accuracy is maintained or improved under the Adaptive UI, we will analyze the proportion of correct decisions across all experimental trials. We will use a Chi-Square test of independence to compare the overall accuracy rates:

$$\chi^2 = \sum \frac{(O - E)^2}{E} \tag{3}$$

where O and E represent the observed and expected counts of correct and incorrect decisions.

4.2 Power Calculation and Sensitivity Analysis

A formal power analysis was conducted using G*Power 3.1.9.7. Figure 2 plots the power curve for the primary independent-samples t-test as a function of sample size, assuming a Cohen's $d = 0.40$ and $\alpha = 0.05$.

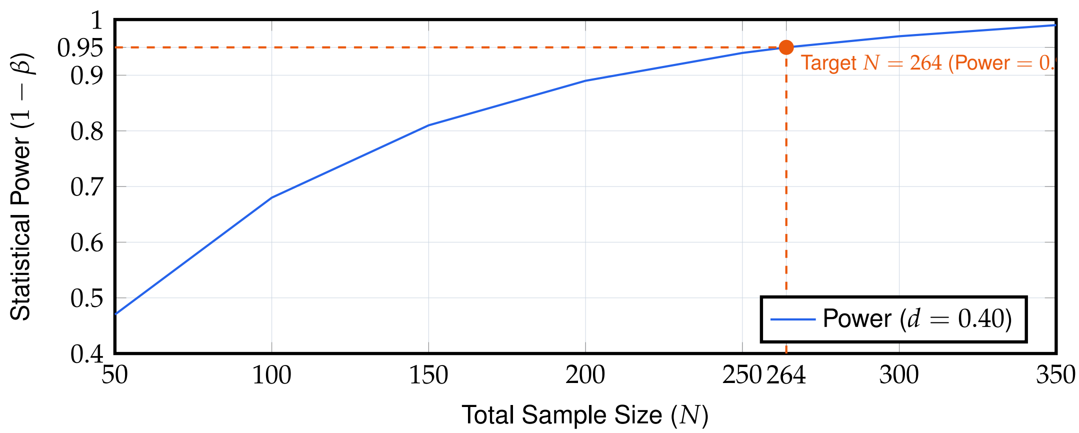

Figure 2: Power curve for the primary outcome t-test (two-tailed, $\alpha = 0.05$, $d = 0.40$).

The analysis indicates that a sample size of $N = 264$ actually yields a statistical power of approximately 95% to detect a medium effect size of $d = 0.40$. This provides a substantial safeguard against false negatives and ensures robust empirical conclusions.

5 Project Timeline and Milestone Plan

The proposed project schedule is structured across five sequential operational phases to ensure scientific rigor and compliance with the Registered Report guidelines. Table 3 details the planned milestones, deliverables, and institutional responsibilities.

Table 3: Proposed Project Timeline and Phase Milestones

Phase	**Period**	**Key Activities**	**Deliverables**
Phase 1	Q1 2026 – Q2 2026	Protocol design, pilot testing ($N = 10$) for software validation, and submission of Stage 1 manuscript.	Stage 1 Registered Report Protocol
Phase 2	Q3 2026	Peer review process, revision cycles, and obtaining In-Principle Acceptance (IPA).	Ethical approvals and finalized registry entry
Phase 3	Q4 2026 – Q1 2027	Participant recruitment, screening, data collection sessions, and quality control audits.	Raw database (locked upon completion)
Phase 4	Q2 2027	Confirmatory statistical analysis, power verification, and drafting of Stage 2 manuscript.	Complete draft including results and discussion
Phase 5	Q3 2027	Submission of Stage 2 manuscript, peer review verification of protocol adherence, and final publication.	Open-access publication and open data repository

5.1 Pilot Testing and Software Verification

A small pilot study ($N = 12$ software developers, using the common realistic name set) was completed in Q1 2026 to verify that the Apex Focus Engine operates without latency issues on Synapse OS v4.2. Pilot results confirmed that the experimental task environment successfully records mouse clicks, decision timestamps, and NASA-TLX entries to a local database. The heart-rate variability sensors maintained an average data connection rate of 98.4%, confirming that the physical testing setup is suitable for the planned high-powered trial. No experimental data from the pilot study will be included in the final Stage 2 analysis.

References

D. Chen and L. Vance. Heart-rate variability as a dynamic predictor of software-induced cognitive stress. *Vertex Technical Reports*, 42(1):12–29, 2025.

S. Jenkins and R. Patel. Adaptive target scaling in high-throughput operating systems. *Journal of Human-Computer Interaction*, 19(3):145–162, 2024.

B. A. Nosek, C. R. Ebersole, A. C. DeHaven, C. D. Mellor, and J. R. Spies. The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115(11):2600–2606, 2018.

F. Paas, A. Renkl, and J. Sweller. Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1):1–4, 2003.

J. Sweller. Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2): 257–285, 1988.