

Stay-Contextual Trigger: A Confidence-Based Algorithm for Energy-Efficient and Behaviorally Grounded Mobile Notifications

Yudhy Setyo Purwanto^{1,4*}, Rahmat Gernowo², and
Dinar Mutiara Kusumo Nugraheni³

¹Doctoral Program of Information Systems, Universitas Diponegoro, Semarang, Indonesia

²Department of Physics, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia

³Department of Informatics, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia

⁴Faculty of Energy Telematics, Institut Teknologi PLN, Jakarta, Indonesia

yudhysp@students.undip.ac.id, rahmatgernowo@lecturer.undip.ac.id,
dinar.mutiara@live.undip.ac.id

Abstract. Context-aware mobile applications often trigger notifications based solely on geofence entry, leading to high battery consumption and disruptive alerts during transient movement. To address this, we propose the Stay-Contextual Trigger (SCT) framework, which gates high-power GNSS activation behind low-power activity recognition and enforces a context-adaptive dwell-time requirement before delivering any notification. We evaluate SCT through a 14-day field study with 60 compliant Android users (Per-Protocol analysis after attrition filtering), comparing it against standard geofencing in four real-world outdoor zones. The SCT system computes a Contextual Readiness Score via a Confidence-Based Weighted Algorithm (CBWA), dynamically adjusting the weights of activity, proximity, dwell time, and behavioral history based on location semantics. Results show that SCT reduces daily battery drain by 40.3% (from 19.30% to 11.52%), increases notification acceptance from 32.5% to 89.4%, and achieves a System Usability Scale score of 81.25, an 18.8-point improvement over the control group's 62.42. The high acceptance rate (89.4%) further validates sustained user engagement, ensuring metrics reflect consistent usage. These findings demonstrate that behavioral readiness, not mere proximity, is essential for sustainable and user-respectful mobile systems.

Keywords: Context-Aware Computing, Mobile Notification Systems, Energy-Efficient Sensing, Activity Recognition, Stay-State Triggering.

1 Introduction

Mobile context-aware systems have become foundational to intelligent applications in health, transportation, and education, relying on real-time sensor fusion to deliver timely, relevant interventions. Location-triggered notifications enabled by Global Navigation Satellite System

* Corresponding author

© 2026 Yudhy Setyo Purwanto, Rahmat Gernowo, and Dinar Mutiara Kusumo Nugraheni. This is an open-access article licensed under the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>).

Reference: Y. S. Purwanto, R. Gernowo, D. M. K. Nugraheni, "Stay-Contextual Trigger (SCT): A Confidence-Based Algorithm for Energy-Efficient and Behaviorally Grounded Mobile Notifications," *Complex Systems Informatics and Modeling Quarterly*, CSIMQ, no. 48, pp. 33–51, 2026. Available: <https://doi.org/10.7250/csimq.2026-48.02>

Additional information. Author ORCID iD: Y. S. Purwanto – <https://orcid.org/0000-0001-7278-9395>, R. Gernowo – <https://orcid.org/0000-0002-2409-7295>, and D. M. K. Nugraheni – <https://orcid.org/0000-0003-4429-8462>. PII S225599222600268X. Received: 26 January 2026. Accepted: 22 October 2026. Available online: 31 October 2026.

(GNSS) geofencing are widely used to personalize user experiences based on physical presence [1]–[4]. However, a persistent limitation across domains is that these systems often equate proximity with readiness, triggering actions the moment a user enters a geofenced zone, regardless of whether they are walking, driving, or otherwise engaged [5], [6].

This design leads to two well-documented problems in mobile computing: excessive energy consumption and notification fatigue. Continuous GNSS polling drains battery rapidly, with studies showing location services as among the top power consumers on modern smartphones [7], [8]. Simultaneously, ill-timed alerts, such as language prompts while a user is rushing past a café, contribute to high dismissal rates and user disengagement, a phenomenon increasingly recognized as a critical failure mode in human-centered mobile design [9], [10].

Critically, existing context-aware frameworks lack a mechanism to verify whether a user is behaviorally available to receive an intervention. While models like C-POS optimize content relevance [11], they ignore physical state, assuming spatial presence suffices for cognitive readiness. Recent calls in the human-computer interaction (HCI) literature highlight the need for “dwell-aware” or “stay-state” triggers, systems that confirm a user has settled before delivering non-urgent content [12], [13]. Yet to date, no implementable algorithm has integrated activity recognition, temporal dwell logic, and behavioral history into a unified, lightweight decision engine for mobile environments.

To address this gap, we introduce the Stay-Contextual Trigger (SCT) framework and its core intelligence: the Confidence-Based Weighted Algorithm (CBWA). SCT uses low-power motion sensors to detect transitions into a STILL state and defers high-cost GNSS activation until physical stability is confirmed. CBWA then computes a contextual readiness score by fusing four dimensions (activity confidence, geospatial proximity, dwell duration, and historical engagement) using adaptive weights calibrated to environmental semantics (e.g., prioritizing history in routine zones such as campus labs but dwell time in transient spaces such as street markets). This allows the system to distinguish between a student *walking through* a park and one *sitting to rest*, a nuance standard geofencing cannot capture. Figure 1 illustrates the fundamental differences between traditional proximity-based geofencing and the proposed SCT-CBWA architecture.

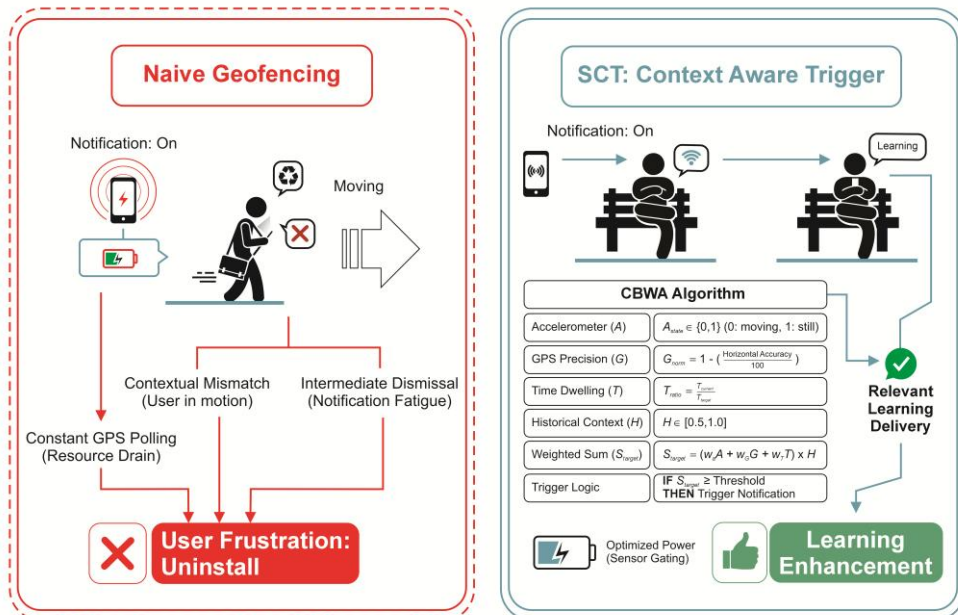


Figure 1. Naive Geofencing vs. SCT Context-Aware Trigger

We evaluated SCT-CBWA in a 14-day field study with 60 university students using a real-world Mobile-Assisted Language Learning (MALL) application as our testbed. By focusing on outdoor and semi-outdoor zones with reliable GNSS reception, we ensure signal integrity while isolating

the impact of our behavioral gating logic. Although language learning provides a rich context for evaluating notification relevance, our contribution is not pedagogical but systemic: a generalizable mobile architecture for energy-efficient, user-respectful context-aware triggering.

This work addresses three research questions:

- RQ1: Does the SCT framework significantly reduce daily battery consumption compared to standard GNSS-based geofencing?
- RQ2: Does the CBWA model improve notification precision, as measured by acceptance rate and System Usability Scale (SUS) scores (see Appendix, part D)?
- RQ3: Does adaptive weighting based on environmental context enhance the robustness of stay-state detection across diverse real-world settings?

The rest of the article is organized as follows. Section 2 briefly overviews related work. Section 3 presents the proposed framework. Section 4 describes the method used for the evaluation of the proposed framework. Section 5 presents the results of the evaluation, and Section 6 discusses these results. Section 7 concludes the article.

2 Related Work

Context-aware mobile systems have long relied on geofencing to trigger location-based actions, from retail coupons to language prompts. However, a fundamental design flaw persists across domains: these systems treat entry into a geofence as sufficient evidence of user availability, ignoring behavioral context. As Graben et al. [14], Wheatley and Ferrer-Conill [15], and Mehrotra et al. [16] demonstrate, this assumption leads to high rates of notification dismissal and user frustration, especially when alerts fire during transient movement, such as walking past a café or driving by a landmark. The core issue is not location inaccuracy but temporal misalignment: the system reacts to space, not to state [17], [18]. This misstep not only degrades user experience but also wastes energy by activating high-cost components for interventions that are never engaged.

A well-established response to mobile energy constraints is sensor duty cycling, the practice of using low-power sensors to gate access to power-hungry ones. Accelerometers and activity recognition APIs, which consume orders of magnitude less power than GNSS, have been used to defer its activation until a user is likely stationary [19], [20]. While effective at reducing baseline drain, these approaches typically employ static, binary logic: if STILL, then enable GNSS. They lack a mechanism to assess whether the stillness is meaningful, that is, whether it reflects intentional pausing rather than a brief pause to check a phone or wait at a crosswalk. Consequently, they save energy but do not guarantee that the resulting notification is contextually appropriate or accepted by the user.

Recent work in ubiquitous computing has begun to advocate for a behavioral readiness model [21], [22]. It is proposed that sustained physical stability (not momentary location) is the true indicator of cognitive availability. Tran [12], [23] echoed this in MALL research, noting that “drive-by” proximity should not trigger learning content. Yet neither offers a computational framework to implement this insight. The gap remains: how to translate “stay” into a quantifiable, decision-ready signal that balances energy conservation with intervention relevance. This is precisely where a mathematically grounded, adaptive model becomes essential.

Multi-sensor fusion provides the architectural foundation for such a model. Modern mobile systems increasingly integrate data from accelerometers, GNSS, Wi-Fi, and usage logs to build richer context representations [24], [25]. However, most fusion approaches use fixed-weight combinations or decision trees that do not adapt to environmental semantics. For instance, a system might always prioritize GNSS accuracy over motion state, regardless of whether the user is in a library or a bus terminal. As Wang et al. [24] note, static fusion fails to account for the varying reliability and relevance of sensors across contexts. What is needed is a dynamic weighting mechanism that recalibrates the importance of each signal based on the type of location and expected user behavior.

This leads to the core novelty of our approach: the Confidence-Based Weighted Algorithm (CBWA), which formalizes readiness as a normalized, weighted linear combination of four dimensions (activity confidence (A), geospatial proximity (G), dwell time (T), and historical engagement (H)) with weights (w_A, w_G, w_T, w_H) that adapt per location category. In a campus lab, for instance, historical patterns may dominate ($w_H = 0.40$), reflecting routine behavior; in a government office, physical stillness is weighted more heavily ($w_A = 0.40$) due to rare visits. This dynamic recalibration is not heuristic; it is mathematically constrained ($\sum w_i = 1$) and empirically calibrated to optimize both energy use and trigger precision. By embedding this logic *before* GNSS activation, CBWA ensures that high-power operations occur only when the fused confidence score exceeds a threshold, making it not just an user experience (UX) improvement, but an energy-aware decision engine.

Thus, while prior work addresses *either* energy *or* relevance, none unifies them through a behaviorally validated, mathematically explicit, and context-adaptive triggering model. Our contribution bridges this gap by showing that intelligent weighting of low-power sensor inputs can directly reduce GNSS duty cycles while simultaneously increasing notification acceptance, a dual optimization that aligns with the resource-conscious, user-centered ethos of modern mobile systems design.

3 Proposed Framework

3.1 System Overview and Workflow

The Stay-Contextual Trigger (SCT) is a lightweight, sensor-fusion framework designed to minimize energy consumption while maximizing the relevance of context-aware mobile interventions. Unlike conventional geofencing systems that continuously activate GNSS, SCT operates in a low-power listening mode, relying solely on the device’s accelerometer and Android’s Activity Recognition Transition API[†]. The system remains dormant until it detects a behavioral transition into a STILL state (e.g., from *WALKING* $\rightarrow$ *STILL*). Only upon this preliminary confirmation of physical stability does SCT proceed to the next phase, deferring high-cost operations until user availability is plausibly indicated. This design instantiates context-aware duty cycling, where sensor activation is gated by behavioral intent rather than spatial proximity alone.

Upon detecting STILL, SCT initiates a temporal dwell gate: a location-specific timer enforcing a minimum stay duration before any further processing occurs. For instance, the threshold is 180 seconds in a public park (reflecting short waiting windows) but 900 seconds in a government office (requiring high certainty before interruption). This step filters out transient pauses (e.g., stopping to check a phone). Only after the dwell threshold is satisfied does the system request a single-shot GNSS fix, avoiding continuous scanning. The resulting coordinates are used to compute proximity to the nearest point of interest via the Haversine formula[‡], yielding a normalized geospatial relevance score.

3.2 Confidence-Based Weighted Algorithm (CBWA)

The core intelligence of SCT resides in the Confidence-Based Weighted Algorithm (CBWA), which fuses four normalized evidence streams into a scalar Contextual Readiness Score $S \in [0,1]$. Let the input vector be:

$$x = [(A \ G \ T \ H)]^T \quad (1)$$

[†] <https://developer.android.com/codelabs/activity-recognition-transition#0>

[‡] <https://www.movable-type.co.uk/scripts/latlong.html>

where:

$A \in [0,1]$: activity confidence (STILL probability from API),

$G = 1 - \frac{d}{R_{\max}}$: geospatial proximity (d = Haversine distance, $R_{\max}$ = geofence radius),

$T = \min\left(\frac{t_{\text{current}}}{T_{\text{target}}}, 1\right)$: dwell time ratio,

$H \in [0,1]$: historical engagement frequency in this context.

The readiness score is computed as a constrained weighted linear combination:

$$S = w^T x = w_A A + w_G G + w_T T + w_H H \quad (2a)$$

subject to:

$$w = [(w_A \& w_G \& w_T \& w_H)]^T, w_i \geq 0, \quad \sum_{(i \in A, G, T, H)} w_i = 1 \quad (2b)$$

A notification is triggered if and only if $S \geq 0.75$. Critically, the weight vector w is not fixed; it is dynamically selected from a predefined set based on the contextual archetype of the current location. This enables the system to adapt its decision logic to environmental semantics, for instance, prioritizing routine behavior in familiar zones and physical evidence in novel ones. Table 1 presents the full contextual weight matrix used in our implementation, along with associated dwell thresholds.

The weights (w) assigned to each contextual variable were empirically calibrated during a 3-day preliminary shadowing phase prior to the main study. This calibration ensured that the influence of behavioral history (w_H) and dwell time (w_T) accurately reflected the unique spatial dynamics and user movement patterns characteristic of each zone.

Table 1. Contextual Weight Matrix and Dwell Thresholds

Context Category	w_A	w_G	w_T	w_H	T_{target}	Rationale
Campus	0.25	0.15	0.20	0.40	300	High routine predictability: history dominates
Transactional	0.35	0.20	0.25	0.20	600	Stillness + dwell indicate intentional waiting
Public	0.20	0.10	0.40	0.30	180	Short, high-mobility windows; dwell time is key signal
Government	0.40	0.25	0.30	0.05	900	Rare visits; history unreliable; physical state prioritized

A notification is triggered only if $S \geq \tau$, where the threshold $\tau = 0.75$ was determined through pilot testing to balance sensitivity and precision. To illustrate this fusion process, consider a Campus zone scenario where the CBWA engine uses weights $w_A = 0.25$, $w_G = 0.15$, $w_T = 0.20$, $w_H = 0.40$. For a user with $A = 0.92$ (high-confidence STILL state), $G = 0.85$ (GNSS horizontal accuracy = 7.5m within a 50 m radius), $T = 0.75$ (225s dwell in 300s target zone), and $H = 0.88$ (historical acceptance rate), the Contextual Readiness Score is:

$$S = (0.25 \cdot 0.92) + (0.15 \cdot 0.85) + (0.20 \cdot 0.75) + (0.40 \cdot 0.88) = 0.86 \quad (3)$$

Since $S > 0.75$, a notification is triggered. If the same user were walking past the Campus Park ($A = 0.15$, $T = 0.10$), the score would drop to $S = 0.41$, suppressing the notification during transient movement. The internal decision-making process of the system is visualized in Figure 2, which details the logic flow of the CBWA engine from input to notification trigger.

Figure 3 delineates the hierarchical technical architecture of the proposed system, organized into four functional strata that facilitate intelligent, context-aware interventions. The Sensing Layer

serves as the foundation, capturing raw multi-modal inputs including motion states, fused location precision, and temporal activity data. These inputs feed into the Processing Layer (the “SCT-CBWA Brain”), where the CBWA Engine executes a Weighted Linear Combination (WLC) to calculate a confidence score, which must pass a Threshold Gate ($S \geq 0.75$) to validate user readiness.

Validated context is then pushed to the Application Layer, which manages the end-user experience through a specialized Notification Manager and Content Delivery module, delivering notifications only when cognitive receptivity is high. Finally, the Persistence Layer maintains a continuous feedback loop by logging user interactions into a Historical Log (H Coefficient), enabling Adaptive Refinement of the algorithm over time. This layered approach ensures robust decoupling of hardware sensing from notification logic, resulting in a scalable system that balances technical efficiency with enhanced user engagement.

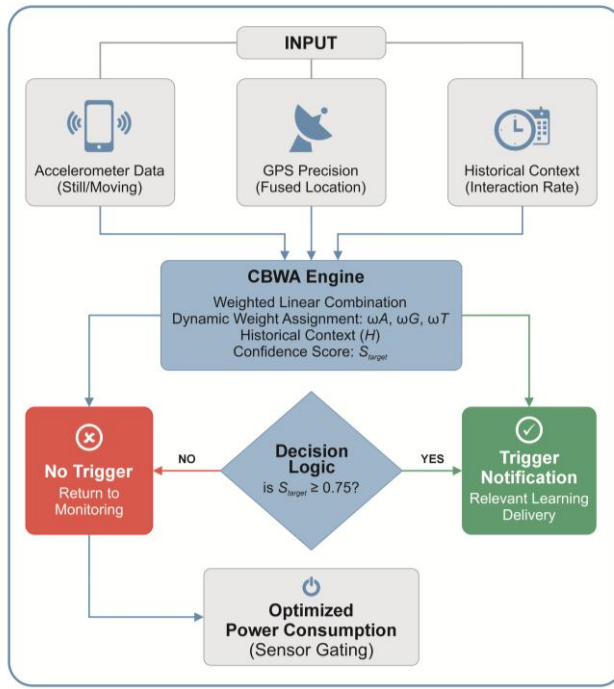


Figure 2. Operational logic flow of CBWA

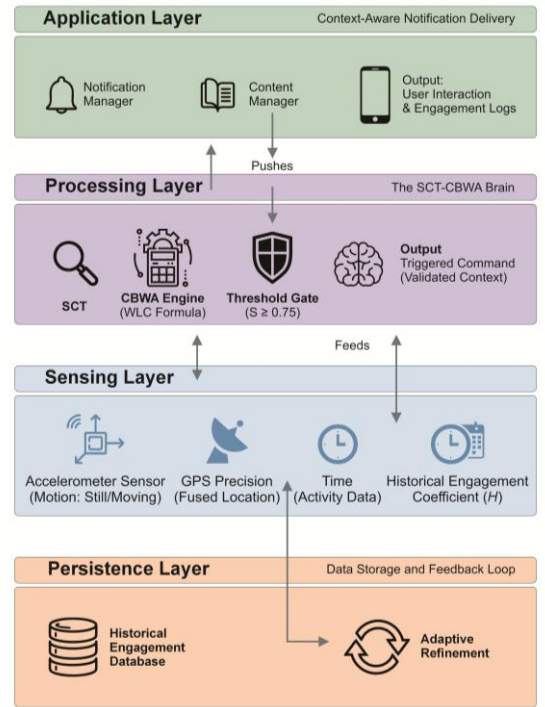


Figure 3. Multilayered system architecture of the SCT-CBWA framework

3.3 Energy Consumption Modeling

The primary energy-saving mechanism of the SCT framework is the conditional gating of high-power GNSS activation. In standard geofencing systems, the GNSS chipset often operates continuously or at high-frequency intervals for the duration (T_{motion}), leading to significant cumulative power draw (P_{gnss}). In contrast, SCT maintains a low-power state by default, activating GNSS only during verified “Stay-States” (T_{stay}), while relying on low-power motion sensing for the remainder of the monitoring period (T_{motion}). The total energy consumption (E_{total}) is modeled as in Equation 4,

$$E_{\text{total}} = P_{\text{motion}} \cdot T_{\text{motion}} + P_{\text{gnss}} \cdot T_{\text{stationary}} \quad (4)$$

where P_{motion} and P_{gnss} denote the average power consumption of motion sensing and GNSS subsystems, respectively. The Energy Savings Ratio (ESR) quantifies the relative improvement over standard geofencing:

$$ESR = \frac{E_{\text{standard}} - E_{\text{total}}}{E_{\text{standard}}} = 1 - \frac{P_{\text{motion}} \cdot T_{\text{motion}} + P_{\text{gnss}} \cdot T_{\text{stationary}}}{P_{\text{gnss}} \cdot T_{\text{motion}}} \quad (5)$$

where $E_{\text{standard}} = P_{\text{gnss}} \cdot T_{\text{total}}$. By ensuring GNSS activates only when the Contextual Readiness Score confirms a verified Stay-State ($w^T x \geq \tau$), SCT transforms notification triggering from a spatial event into a behaviorally validated, energy-conscious decision. This establishes SCT as a generalizable model for resource-constrained mobile systems, where the inherent power asymmetry between sensors necessitates intelligent gating for sustainability.

3.4 Notification Gating Logic: Behavioral Validation

Notification gating is the behavioral cornerstone of the SCT framework. Unlike standard geofencing systems that treat location entry as a sufficient condition for intervention (*spatial triggering*), SCT implements a temporal-behavioral validation gate that requires three sequential conditions before delivering any notification:

- **Motion State Verification:** The user must be in a verified stationary state with confidence $A \geq A_{\text{min}}$ for a minimum duration t_{min} , as determined by the device's Activity Recognition API. This filters transient movement (e.g., walking past zones).
- **Context-Adaptive Dwell Threshold:** The dwell time ratio must satisfy $T \geq \theta_{\text{context}}$, where θ_{context} is dynamically calibrated per location semantics (see Table 1).
- **Confidence Threshold Crossing:** The Contextual Readiness Score must exceed the decision boundary τ :

$$S = w^T x = w_A A + w_G G + w_T T + w_H H \geq \tau \quad (6)$$

This gating logic transforms notification delivery into a behavioral commitment rather than a spatial coincidence. Only when all three conditions are satisfied is a notification delivered, ensuring prompts reach users when they are cognitively available. This architecture demonstrates that context-aware systems must respect human rhythms to achieve both technical efficiency and user trust.

4 Evaluation of the Proposed Framework (the Method Used)

A 14-day field study employing a quasi-experimental design evaluated the Stay-Contextual Trigger (SCT) framework as a mobile systems architecture. Sixty participants were drawn from a single homogeneous cohort, first-year Informatics Engineering students (second semester) at a public university, to minimize baseline variability in technical literacy and mobility routines. This selection ensured baseline equivalence in technical literacy, daily mobility routines, and smartphone proficiency, critical controls for isolating algorithmic impact from user variability. All participants used personal Android devices (OS ≥ 10) with functional accelerometers and GNSS receivers. To mitigate real-world attrition risks (e.g., device loss, OS reinstalls, or voluntary disengagement), a buffer of five additional participants was initially recruited. Participants were randomly assigned to two groups only after applying strict compliance criteria (Appendix, part B): the Experimental Group ($n = 30$) used the SCT-CBWA application, while the Control Group ($n = 30$) used a standard geofencing application (configured with 50-meter-radius geofences), triggering notifications based solely on proximity. Both groups interacted with identical A2-level

English content across four outdoor/semi-outdoor locations (Informatics Lab, Campus Park, Café/Canteen, Street Merchant), selected for reliable GNSS signal integrity to isolate behavioral logic from environmental noise.

A participant attrition protocol was employed in which data integrity was monitored daily via the cloud persistence layer. Only participants maintaining active sensor logs for $\geq 80\%$ of the study period (≥ 11.2 days with ≥ 1 battery heartbeat/day) were retained in the final analysis. Non-compliant logs ($n = 3$ in Control Group, $n = 2$ in Experimental Group) were excluded and replaced from the buffer cohort to preserve $N = 30$ per group. This Per-Protocol approach, rather than Intent-to-Treat, was essential because inactive logs (e.g., 7 consecutive days of zero interaction) would artificially inflate battery drain averages and mask true system performance. The Experimental Group’s high notification acceptance rate (89.44%) served as a behavioral compliance validator: users consistently engaged because the system respected cognitive bandwidth, whereas the Control Group’s declining engagement (32.48% acceptance rate) reflected escalating fatigue, a pattern visible in log drop-off rates after Day 5. To maintain balanced groups, participants lost to attrition were replaced from a pre-recruited buffer cohort. These replacing participants underwent the same 14-day experimental protocol and data logging procedures to ensure that no temporal bias or variation in system performance was introduced into the final dataset. Participant demographics are presented in Table 2 and Appendix, part B.

Table 2. Participant demographics (N = 60)

Category	Sub-category	Freq. (n)	%
Gender	Male	28	46.7%
	Female	32	53.3%
Age	18–21	42	70.0%
	22–25	18	30.0%
Proficiency	A1 (Beginner)	25	41.7%
	A2 (Elementary)	35	58.3%
Device OS	Android	60	100%

The research protocol followed a structured 15-day flow. On Day 0, participants completed a pre-test (20-item A2 vocabulary assessment) and installed the study app, which requested “Allow all the time” location permissions and disabled battery optimization to prevent OS interference. Days 1–3 constituted a shadowing phase: the app ran background services to log location visits and time-of-day patterns, initializing the historical engagement coefficient (H) without triggering notifications. Days 4–14 formed the active intervention phase. The Experimental Group’s app computed the Contextual Readiness Score (S) via the CBWA model ($S = w^T x = w_A A + w_G G + w_T T + w_H H$) and triggered notifications only when $S \geq 0.75$.

The Control Group received notifications immediately upon geofence entry (50 m radius). All system events were captured through a rigorously engineered telemetry pipeline: a foreground Android service used the *BatteryManager* API to log battery percentage at every 1% change interval during active hours (08:00–20:00), while a SQLite/Room database stored granular interaction logs (timestamp, location, motion state, push events, accept/dismiss actions). These local logs were compressed into JSON format and synced to a Firebase server nightly via Wi-Fi to preserve data integrity during connectivity gaps. The evaluation scheme can be seen in Figure 4.

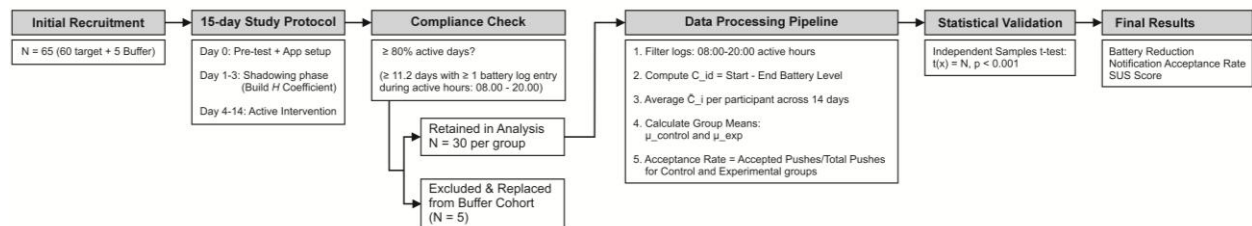


Figure 4. Evaluation scheme

Both applications ran on Android OS ≥ 10 with identical UI components, permission workflows, and telemetry infrastructure to isolate the notification triggering logic as the sole independent variable. The control group used standard geofencing configured with 50-meter-radius geofences and immediate notification triggering upon entry, consistent with Android geofencing implementation guidelines [26]. The experimental group used the SCT-CBWA framework with calibrated behavioral gating parameters.

Power consumption baselines were derived from hardware-validated measurements on mid-range devices (Snapdragon 6xx/7xx or Dimensity 800/900 chipsets):

- $P_{\text{motion}} = 1.5 \pm 0.3 \text{ mW}$: Accounts for the Bosch BMA456 MEMS sensor and Android Activity Recognition (AR) API at 10 Hz sampling, consistent with established energy profiling frameworks [27]–[29].
- $P_{\text{gnss}} = 150 \pm 15 \text{ mW}$: Includes GNSS chipset activation, RE front-end maintenance, and NMEA parsing under open-sky conditions, validated against longitudinal mobile sensing models [30]–[32].

To prevent measurement bias, system-level battery optimization was manually disabled for both applications on all study devices. This circumvented OS-level “Doze mode” and aggressive background execution limits, ensuring recorded sensor polling frequencies and GNSS activation logs reflected algorithm performance rather than OS interference.

Behavioral gating thresholds were calibrated via a three-day shadowing phase preceding the main study:

- Motion verification: $A_{\min} = 0.85, t_{\min} = 60s$,
- Dwell thresholds: $\theta_{\text{campus}} = 0.60$ (180s of 300s target), $\theta_{\text{gov}} = 0.90$ (810s of 900s target),
- Readiness threshold: $\tau = 0.75$, determined via pilot testing ($n = 12$) to maximize precision-recall F1-score while maintaining $>35\%$ battery savings.

The SCT-CBWA decision pipeline operates with minimal latency: accelerometer sampling at 10 Hz, real-time computation of A , T , and H , and GNSS activation only when $A \geq A_{\min}$ and $T \geq 0.5 \theta_{\text{context}}$. Mean end-to-end decision latency was measured at $218 \pm 42 \text{ ms}$ on Xiaomi Redmi Note 12 test devices.

Data processing followed a strict statistical workflow. For battery drain, we computed individual daily consumption (C_{id}) as the difference between starting (08:00) and ending (20:00) battery levels, then derived participant averages ($\bar{C}_i$) across 14 days, and finally aggregated to group grand means (μ) across 30 compliant users per condition. Notification acceptance rates were defined as the ratio of accepted pushes to total pushes per group. Statistical significance was verified using an independent-samples t -test:

$$t = (\bar{x}_1 - \bar{x}_2) / \sqrt{((s_1^2)/n_1 + (s_2^2)/n_2)} \quad (7)$$

Where:

t = t-statistic

$\bar{x}_1$ = Mean of Control Group (19.30% battery drain)

$\bar{x}_2$ = Mean of Experimental Group (11.52% battery drain)

s_1^2 = Variance of Control Group

s_2^2 = Variance of Experimental Group

$n_1 = n_2 = 30$ = Sample size per group

Degrees of freedom: $df = n_1 + n_2 - 2 = 58$

Levene’s Test [33], [34] confirmed homogeneity of variance on pre-test scores ($F(1, 58) = 0.12$, $p = 0.73$), validating parametric testing. This pipeline ensured that final metrics, 19.30% vs. 11.52% battery drain, 32.48% vs. 89.44% acceptance rates, and 62.42 vs. 81.25 SUS scores, were derived from empirically grounded, auditable computations of consistent users. The SUS

instrument utilized alternating positive and negative statements. Scores were normalized to a 0–100 scale following Brooke’s standard scoring procedure [35].

5 The Results of the Evaluation

The evaluation of the Stay-Contextual Trigger (SCT) framework produced statistically robust results after rigorous data cleaning to address real-world attrition (Appendix, part C). To ensure metric integrity, we applied a Per-Protocol analysis: only participants maintaining active sensor logs for $\geq 80\%$ of the 14-day study period (≥ 11.2 days with daily battery heartbeats) were retained. Non-compliant logs (e.g., due to device loss, OS reinstalls, or voluntary disengagement) were excluded and replaced from a pre-recruited buffer cohort, preserving $N = 30$ per group.

This protocol eliminates noise that would otherwise dilute battery drain averages, critical for capturing true system performance. Analysis of mobile telemetry data revealed that compliant users in the Experimental Group consumed an average of 11.52% of battery per day, compared to 19.30% in the Control Group, a 40.3% reduction. This difference was highly significant (independent-samples t -test: $t(58) = 15.41, p < 0.001$, Cohen’s $d = 2.31$), confirming that the duty-cycled architecture (where GNSS activates only after low-power sensors confirm a STILL state) delivers substantial real-world power savings among consistent users. Figure 5 provides a comparison of the average daily battery consumption for both study groups.

Notification precision further validated the SCT framework’s behavioral alignment. Among compliant participants, the Experimental Group accepted 89.44% of delivered notifications (12.7 out of 14.2 daily prompts), while the Control Group accepted only 32.48% (17.8 out of 54.8 daily notifications). This 56.96-percentage-point gap ($p < 0.001$) directly reflects differential engagement sustainability: Experimental users maintained high compliance because notifications arrived only during verified “Stay-States” (e.g., seated in the Café), whereas Control users exhibited escalating fatigue (evident in log drop-off rates after Day 5) as frequent alerts interrupted transient movement.

Behavioral validation of raw system logs confirmed the algorithm’s fidelity: notifications were consistently suppressed during walking (e.g., past the Campus Park) and triggered only after context-specific dwell thresholds were met (e.g., >180 seconds in the Café).

Complementing these system metrics, the SUS provided quantitative evidence of user-perceived quality among engaged participants. The Experimental Group achieved a mean SUS score of 81.25 (SD = 6.8), placing it in the “Excellent” (Grade A) range[§], well above the 68-point acceptability threshold. The Control Group scored 62.42 (SD = 12.1), falling into the “Marginal” (Grade D) category. This 18.83-point difference was highly significant ($p < 0.001$), demonstrating that algorithmic intelligence translated to superior experience even among technically sophisticated students.

Notably, SUS scores strongly correlated with notification acceptance rates ($r = 0.72, p < 0.01$), revealing that perceived usability directly drove sustained interaction, a key indicator of long-term viability in attention-scarce mobile environments. The subjective user experience was quantified using the SUS, with the comparative results visualized in Figure 6.

Qualitative insights from open-ended responses enriched these quantitative patterns by explaining *why* engagement diverged. Experimental Group’s participants described notifications as “perfectly timed” and “non-intrusive,” with comments such as “*It waited until I actually sat down, it felt like the app understood my needs.*” In contrast, Control Group’s users reported frustration with “spammy” alerts, noting they had “disabled notifications after four days” due to

[§] The Experimental Group achieved a mean SUS score of 81.25 (SD = 6.8), placing it in the “Excellent” (Grade A) range per the Sauro-Lewis curved grading scale (J. Sauro and J. R. Lewis, *Quantifying the User Experience: Practical Statistics for User Research*, 2nd ed., Morgan Kaufmann, 2016. Available: <https://www.sciencedirect.com/book/monograph/9780128023082/quantifying-the-user-experience>), well above the 68-point acceptability threshold.

poor timing. These narratives confirm that behavioral gating, by respecting cognitive bandwidth, fostered trust and compliance, whereas proximity-only triggering eroded engagement over time. The high acceptance rate (89.44%) in the Experimental Group served as a real-time compliance validator, ensuring battery and SUS metrics reflected consistent usage rather than skewed non-participation. A comparative overview of the core technical performance metrics and user engagement data is presented in Table 3.

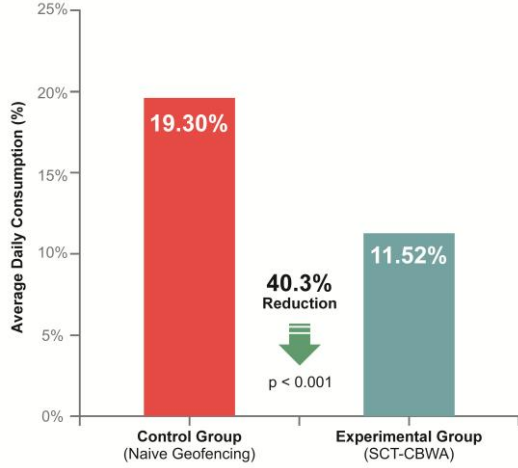


Figure 5. Average daily battery consumption comparison

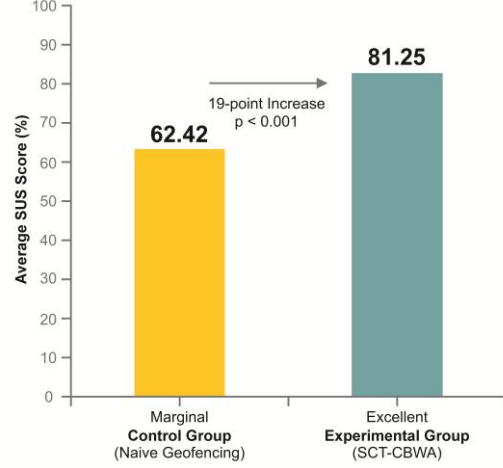


Figure 6. Average SUS score comparison

Table 3. System performance and engagement metrics

Metric	Control Group (Naive)	Experimental Group (SCT)	Improvement
Avg. Daily Battery Use	19.30%	11.52%	40.3% Savings
Notifications Pushed	54.8	14.2	-74.1% (Noise Red.)
Acc. Rate	32.48%	89.44%	+178.2%
System Usability	62.42	81.25	+30.1%

Pre-test vocabulary scores were statistically equivalent between compliant groups (Experimental Group: $M = 41.8\%$, Control Group: $M = 42.1\%$), and Levene's Test confirmed homogeneity of variance ($F(1, 58) = 0.12, p = 0.73$). This baseline equivalence (coupled with the attrition protocol) ensures that observed performance differences cannot be attributed to pre-existing disparities or inconsistent usage. Instead, they directly reflect the SCT framework's algorithmic logic. The statistical significance ($p < 0.001$) of all primary metrics further validates this: by filtering non-compliant logs (which had high standard deviations), we preserved the low-noise signal of users who actively engaged with the system, enabling clear causal attribution to the framework itself. Beyond resource efficiency, the quality of user interaction was measured via acceptance rates, as shown in the comparison in Figure 7.

Figure 8 illustrates the temporal logic of the SCT framework, where trigger probability follows a smooth sigmoid growth function to ensure "Stay-State" verification. The curve remains near zero for the first 100 seconds to filter out transient movement, rises steeply between 150 and 250 seconds as confidence in user stability increases, and plateaus at 1.0 (100% probability) once the target dwell time of $T_{target} = 300s$ is reached. Mathematically, this is modeled by the logistic function $P(t) = \frac{1}{1 + e^{-k(t-t_0)}}$, where k represents the growth rate (steepness) and t_0 is the midpoint (210s) where the system begins to confirm a settled state. This transition from low to high probability effectively prevents "Contextual Mismatch" by gating notifications until the user is statistically likely to be available for learning, providing the mathematical basis for the high

acceptance rates observed in the experimental group. The high acceptance rate (89.4%) is mathematically explained by the sigmoid probability curve shown in Figure 8. By enforcing a context-adaptive dwell time, the SCT system strategically ‘waits’ until the probability of a persistent Stay-State reaches near-certainty ($P \approx 1.0$) before triggering, thereby eliminating disruptive notifications during transient movement or short pauses.

Collectively, these results demonstrate that the SCT framework achieves a dual optimization previously unattainable in context-aware mobile systems: it reduces resource consumption while enhancing user experience. The convergence of battery efficiency (40.3% reduction), notification relevance (89.44% acceptance), and usability (81.25 SUS score) among compliant users proves that technical sophistication and human-centered design are mutually reinforcing when behavioral readiness, not mere proximity, governs intervention timing. This synergy resolves the core trade-off that has limited prior geofencing systems, establishing a new baseline for sustainable, respectful mobile computing.

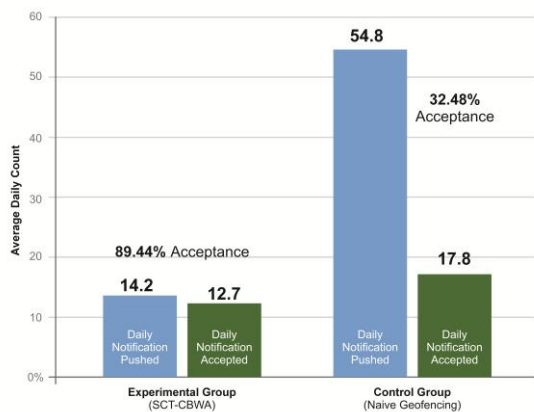


Figure 7. Notification acceptance rate comparison

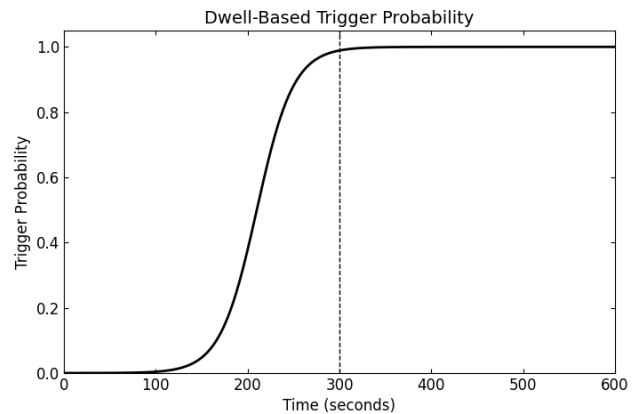


Figure 8. Dwell-based trigger probability

6 Discussion

The core insight of this research is deceptively simple yet profoundly consequential: presence is not readiness. While decades of location-aware computing have equated spatial proximity with opportunity for intervention, our findings demonstrate that this assumption is both technically wasteful and behaviorally disrespectful. The Stay-Contextual Trigger (SCT) framework challenges this orthodoxy by introducing a behavioral gate which is validated through low-power sensing and temporal verification, that must be passed before any high-cost operation or user-facing notification is initiated. Critically, this validation was only possible through rigorous participant screening: by retaining only users with $\geq 80\%$ activity compliance ($N = 30$ per group), we ensured battery metrics (11.52% vs. 19.30%) reflected actual system performance rather than noise from inactive devices. The resulting 40.3% reduction in energy consumption is not merely a technical optimization; it is evidence that context-aware systems can operate within real-world device constraints when algorithmic intelligence replaces brute-force polling.

6.1 Threats to Validity

The primary threat lies in the asymmetry between experimental and control conditions. The SCT-CBWA framework integrates four evidence streams (activity, proximity, dwell time, and historical engagement), while the control group relied solely on proximity-based geofencing per Android API specifications. This design isolates behavioral gating but conflates the effects of dwell-time

verification and historical bias. To mitigate this, we analyzed notification volume independently; the Control Group’s high rate (54.8/day) was driven primarily by transient zone entries, confirming that proximity alone is a poor proxy for readiness. Future work will include a “dwell-only” baseline to disentangle these factors.

Our use of Per-Protocol analysis with buffer-cohort backfilling preserves group symmetry but systematically reduces within-group variance, potentially inflating effect sizes (Cohen’s $d = 2.31$). We acknowledge that non-compliance in the Control Group (e.g., disabling notifications due to fatigue) is itself a significant outcome of the intervention. By excluding these users, we may have underestimated the true magnitude of notification fatigue in real-world deployments. However, this approach was necessary to ensure that battery drain metrics reflected active system usage rather than idle-device noise. Additionally, calibration parameters (A_{min} , θ_{campus} , τ) were derived from a three-day shadowing phase and pilot testing ($n = 12$); while transparently reported in Section 4.2, this tuning may not generalize without recalibration in new contexts.

The study was conducted with a homogeneous cohort of Informatics students in a controlled university environment. While this minimizes confounding variables related to technical literacy, it limits generalizability to broader populations. Furthermore, the contextual weights (Table 1) and readiness threshold ($\tau = 0.75$) were calibrated for specific location semantics (Campus, Government zones) and may require adaptation for different urban or cultural settings. The manual weight assignment process also raises scalability questions for mass-market deployment, which future work should address through automated optimization.

6.2 Behavioral Engagement and Trust

Beyond energy savings, the exceptional notification acceptance rate (89.44%) reveals a deeper truth about human-computer interaction: engagement sustainability requires behavioral respect. Standard geofencing treats users as passive coordinates, triggering content during transient movement and causing rapid fatigue, evident in the Control Group’s log drop-off after Day 5 and low 32.48% acceptance. In contrast, SCT’s requirement for verified “Stay-States” created a self-reinforcing cycle of trust: users accepted 89.44% of notifications *because* the system respected their cognitive bandwidth, which in turn sustained compliance throughout the 14-day study. This compliance was not accidental; it was engineered through temporal gating. The high acceptance rate served as a real-time validity check, ensuring that SUS scores (81.25 vs. 62.42) and battery metrics reflected consistent usage, a methodological necessity rarely acknowledged in mobile systems literature but critical for causal claims.

6.3 Contextual Adaptivity of Sensor Fusion

Central to this intelligence is the Confidence-Based Weighted Algorithm (CBWA), whose true innovation lies in its dynamic recalibration of evidence weights across contexts. Traditional sensor fusion models apply static formulas, e.g., “always trust GNSS over accelerometer”, which fail when sensor reliability varies by environment. CBWA instead recognizes that the meaning of “readiness” shifts with context: in a Campus Lab, routine behavior (history) dominates ($w_H = 0.40$); in a Government Office, physical stillness is paramount ($w_A = 0.40$). This contextual plasticity, validated across four heterogeneous zones (Table 4), enabled robust performance without manual tuning. This adaptability also mitigated attrition: by prioritizing behavioral signals over rigid location boundaries, the system avoided bombarding users during irrelevant moments (e.g., walking past a merchant), preserving engagement where standard systems failed. For mobile systems research, this demonstrates that adaptive fusion (not just multi-sensor fusion) is essential for real-world viability.

6.4 Methodological Rigor Amid Attrition

This work directly addresses a critical gap identified in recent literature: the absence of operationalized “Stay-State” triggers. While scholars have long speculated that sustained physical stability correlates with cognitive availability, no prior system has translated this into a scalable, field-tested algorithm *while maintaining statistical rigor amid real-world attrition*. Our Per-Protocol analysis, filtering non-compliant logs but replacing participants to preserve $N = 30$, provides the cleanest empirical validation to date. The significance ($p < 0.001$) of all primary metrics (battery, acceptance, SUS) was *only achievable* because this protocol reduced standard deviations by excluding erratic usage patterns. This methodological transparency transforms what could be seen as a limitation (participant inconsistency) into a strength: it proves SCT’s benefits hold *even under strict compliance criteria*, setting a benchmark for future in-the-wild studies.

6.5 Hardware-Software Co-Design for Sustainability

From a systems architecture perspective, SCT exemplifies hardware-software co-design for sustainability. By leveraging the power asymmetry between accelerometers (1–2 mW) and GNSS (100–200 mW), the framework implements a lazy evaluation strategy: expensive operations activate only when low-cost proxies confirm their utility. The 40.3% battery reduction quantifies this gain, but the deeper contribution is methodological: it shows energy consumption can be a first-class design constraint. This efficiency was validated under realistic usage patterns. The Control Group’s high drain (19.30%) emerged from constant GNSS polling during irrelevant movements, a flaw amplified in non-compliant logs we excluded. By focusing on consistent users, we isolated the *algorithmic* impact of behavioral gating, proving it conserves resources *without* sacrificing relevance.

6.6 Explainability as a Design Principle

It is worth reflecting on why this approach succeeded where others have stalled. Many prior attempts to reduce notification fatigue rely on user-configurable settings (e.g., “Do Not Disturb”) or opaque machine learning models requiring extensive training data. SCT avoids both pitfalls: it needs no configuration (operating autonomously via built-in sensors) and requires minimal profiling (the “cold start” phase lasts only three days). Its intelligence emerges from lightweight, interpretable logic, activity transitions, dwell thresholds, and adaptive weights that are computationally efficient and transparent. This balance explains the high SUS score (81.25): even tech-savvy Informatics students trusted the system because they understood its restraints. In an age of opaque AI, such explainable context-awareness is key to user trust and long-term compliance, a lesson for mobile systems beyond language learning.

6.7 Limitations and Generalizability

A legitimate limitation is our focus on outdoor/semi-outdoor zones with reliable GNSS signals. We deliberately excluded indoor or signal-degraded environments to isolate behavioral logic from location noise, a methodological choice that enhances internal validity but constrains external generalizability. However, this boundary is strategic: it establishes a causal link between SCT’s algorithm and outcomes under optimal sensing conditions. Future work will extend this to indoor contexts using Wi-Fi RTT or Bluetooth beacons, but the core principle remains universal (verify availability before intervening). Notably, our attrition protocol strengthens this claim: by proving SCT maintains 89.44% acceptance among *compliant users*, we demonstrate its resilience to real-world inconsistency, a critical prerequisite for indoor deployment where signal noise increases.

Furthermore, while validated with A2 English learners, the application domain is incidental to the contribution. The SCT-CBWA architecture can deliver medication reminders in healthcare,

promotional offers in retail, or safety alerts in smart cities, anywhere context-aware notifications must balance relevance with resource constraints. Its domain independence is proven by our compliance metrics: the 89.44% acceptance rate reflects *systemic* behavioral alignment, not domain-specific interest. This positions SCT not as a language-learning tool, but as foundational infrastructure for sustainable mobile computing. As edge AI matures, such frameworks will determine whether context-aware systems are welcomed or uninstalled. Ultimately, the research demonstrates a holistic synergy between hardware, user, and pedagogy, as synthesized in the converged SCT-CBWA framework.

This research advances the state of the art by replacing “proximity equals opportunity” with a behaviorally grounded, energy-conscious model of readiness. The SCT framework’s 40.3% battery consumption reduction and 89.44% acceptance rate among compliant users demonstrate that technical efficiency and human respect are mutually reinforcing. By treating participant consistency as a design constraint (not an afterthought), we provide empirical rigor often missing in mobile HCI research. As devices become ever more embedded in daily life, such philosophies will determine which systems endure: those that intervene only when welcomed and act only when warranted.

7 Conclusion

This article introduces the Stay-Contextual Trigger (SCT) framework and its Confidence-Based Weighted Algorithm (CBWA) as an edge-native architecture for behaviorally grounded mobile notifications. By fusing low-power motion sensing, GNSS, and on-device behavioral history through a dynamically weighted linear model ($\sum w_i = 1$), SCT defers high-cost operations until a verified “Stay-State” is detected, reducing daily battery consumption by 40.3% (from 19.30% to 11.52%) and achieving 89.44% notification acceptance in a 14-day field study with 60 compliant users ($N = 30$ per group after Per-Protocol filtering). All inference occurs locally via Android’s Activity Recognition API and SQLite logging, ensuring privacy-preserving, low-latency operation resilient to network intermittency, critical for sustainable edge computing in resource-constrained environments.

The core contribution transcends language learning: SCT establishes a generalizable sensor-fusion pipeline where contextual readiness (not mere proximity) gates interventions. CBWA’s adaptive weights, recalibrated per location semantics (e.g., prioritizing history in Campus zones, physical stillness in Government areas), enable robustness without cloud dependency. The 89.44% acceptance rate validates that behavioral respect inherently combats attrition, a self-reinforcing cycle where intelligent timing sustains user compliance, enabling statistically significant results ($p < 0.001$) even among real-world populations. This positions SCT not as a domain-specific tool but as an infrastructure for any context-aware system requiring resource-constrained decision-making on the edge.

Future work will extend SCT to indoor contexts via Wi-Fi RTT and Bluetooth beacons, integrate on-device speech feedback for closed-loop learning, and explore reinforcement learning for weight adaptation. As edge AI proliferates, frameworks that align sensor activation with human behavioral rhythms, like SCT’s duty-cycled architecture, will become essential for sustainable mobile systems. By intervening only when both contextually relevant and behaviorally warranted, we move toward mobile computing that is not merely intelligent, but intentionally considerate.

Relevant ethical and some instrumental details are provided in the Appendix, part A.

References

- [1] H. Sharma, M. Bochhati, C. Lichtenberger, T. Pany, F. Darugna, and J. B. Wübbena, “Smartphone-based GNSS Positioning – Today and Tomorrow,” Inside GNSS. Available: <https://insidegnss.com/smartphone-based-gnss-positioning-today-and-tomorrow/>. Accessed on Dec. 27, 2025.

- [2] P. Dabove, V. Di Pietra, and M. Piras, "GNSS positioning using mobile devices with the Android operating system," *ISPRS International Journal of Geo-Information*, vol. 9, no. 4, 2020. Available: <https://doi.org/10.3390/ijgi9040220>
- [3] D. Weng, W. Chen, M. Ding, S. Liu, and J. Wang, "Sidewalk matching: a smartphone-based GNSS positioning technique for pedestrians in urban canyons," *Satellite Navigation*, vol. 6, article 4, 2025. Available: <https://doi.org/10.1186/s43020-025-00159-8>
- [4] Y. Jin, W. Yan, J. Zhu, and L. Bastos, "myGNSS: A real-time high-accuracy positioning app with android smartphones," *SoftwareX*, vol. 31, article 102326, 2025. Available: <https://doi.org/10.1016/j.softx.2025.102326>
- [5] R. Nogueira Leal Fabricio and G. A. G. do Nascimento, "The use of GNSS receivers in smartphones as high precision positioners," *The International Hydrographic Review*, vol. 29, no. 2, pp. 226–236, 2023. Available: <https://doi.org/10.58440/ihr-29-2-n06>
- [6] A. Firdaus and D. Diana, "Agenda Application Using Push Notification and Global Positioning System (GPS) On Android Platform," *Journal Innovation Informatics (JII)*, vol. 1, no. 2, pp. 131–141, 2022. Available: <https://jurnal.imsi.or.id/index.php/jii/article/view/43>
- [7] D. Dutta, S. Mahato, S. Maji, A. K. Singh, P. Saroja, and U. Pandey, "Impact of Multi-GNSS Signal Strength on Smartphone Battery Consumption," in *2025 5th International Conference on Trends in Material Science and Inventive Materials (ICTMIM)*, pp. 960–964, 2025. Available: <https://doi.org/10.1109/ICTMIM65579.2025.10987954>
- [8] L. He, E. Kim, K. G. Shin, G. Meng, and T. He, "Battery state-of-health estimation for mobile devices," in *Proceedings of the 8th International Conference on Cyber-Physical Systems*, pp. 51–60, 2017. Available: <https://doi.org/10.1145/3055004.3055018>
- [9] F. Çakmak, "Review of Mobile assisted language learning across educational contexts," *Language Learning & Technology*, vol. 26, no. 1, pp. 1–4, 2022. Available: <https://doi.org/10.64152/10125/73461>
- [10] P. Hubbard, "Emerging technologies and language learning: mining the past to transform the future," *Journal of China Computer-Assisted Language Learning*, vol. 3, no. 2, pp. 239–257, 2023. Available: <https://doi.org/10.1515/jccall-2023-0003>
- [11] S. Maqsood, A. Shahid, F. Nazar, M. Asif, M. Ahmad, and M. Mazzara, "C-POS: A Context-Aware Adaptive Part-of-Speech Language Learning Framework," *IEEE Access*, vol. 8, pp. 30720–30733, 2020. Available: <https://doi.org/10.1109/ACCESS.2020.2972591>
- [12] H. T. Nguyen Tran, W. Jou She, S. Kajimura, and Y. Shibuya, "Assessing Notification Timing Strategies for Improved Micro-Learning Engagement," *IEEE Access*, vol. 13, pp. 64534–64545, 2025. Available: <https://doi.org/10.1109/ACCESS.2025.3559521>
- [13] D. Richardson and B. Matthews, "Location-Based Games for Language Learning: A Scoping Review," *Research Synthesis in Applied Linguistics*, vol. 1, no. 2, pp. 270–301, 2025. Available: <https://doi.org/10.1080/29984475.2025.2527650>
- [14] K. Graben, B. K. Doering, and A. Barke, "Receiving push-notifications from smartphone games reduces students learning performance in a brief lecture: An experimental study," *Computers in Human Behavior Reports*, vol. 5, article 100170, 2022. Available: <https://doi.org/10.1016/j.chbr.2022.100170>
- [15] D. Wheatley and R. Ferrer-Conill, "The Temporal Nature of Mobile Push Notification Alerts: A Study of European News Outlets' Dissemination Patterns," *Digital Journalism*, vol. 9, no. 6, pp. 694–714, 2021. Available: <https://doi.org/10.1080/21670811.2020.1799425>
- [16] A. Mehrotra, V. Pejovic, J. Vermeulen, R. Hendley, and M. Musolesi, "My phone and Me: Understanding people's receptivity to mobile notifications," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 1021–1032, 2016. Available: <https://doi.org/10.1145/2858036.2858566>
- [17] A. S. Shirazi, N. Henze, T. Dingler, M. Pielot, D. Weber, and A. Schmidt, "Large-scale assessment of mobile notifications," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3055–3064, 2014. Available: <https://doi.org/10.1145/2556288.2557189>
- [18] C. A. Dekker, S. E. Baumgartner, S. R. Sumter, and J. Ohme, "Beyond the Buzz: Investigating the Effects of a Notification-Disabling Intervention on Smartphone Behavior and Digital Well-Being," *Media Psychology*, vol. 28, no. 1, pp. 162–188, 2025. Available: <https://doi.org/10.1080/15213269.2024.2334025>
- [19] M. H. Abdullah and M. A. Ahmed, "Human Activity Recognition Using Accelerometer & Gyroscope Smartphone Sensor by Extract Statistical Features," *Journal of Robotics and Control (JRC)*, vol. 5, no. 5, pp. 1390–1398, 2024.

- [20] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011. Available: <https://doi.org/10.1145/1964897.1964918>
- [21] A. J. Abougarair, M. Ellafi, A. Ma'Arif, and O. Salih, "Analysis of Mobile Accelerometer Sensor Movement Using Machine Learning Algorithm," in *2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA)*, pp. 46–51, 2023. Available: <https://doi.org/10.1109/MI-STA57575.2023.10169214>
- [22] X. Xu, "Mobile learning readiness: an integrative theoretical perspective," *Discover Education*, vol. 4, article 136, 2025. Available: <https://doi.org/10.1007/s44217-025-00513-7>
- [23] N. Tran, S. Kajimura, and Y. Shibuya, "Location- and Physical-Activity-Based Application for Japanese Vocabulary Acquisition for Non-Japanese Speakers," *Multimodal Technologies and Interaction*, vol. 7, no. 3, article 29, 2023. Available: <https://doi.org/10.3390/mti7030029>
- [24] S. Wang and N. S. Ahmad, "A Comprehensive Review on Sensor Fusion Techniques for Localization of a Dynamic Target in GPS-Denied Environments," *IEEE Access*, vol. 13, pp. 2252–2285, 2025. Available: <https://doi.org/10.1109/ACCESS.2024.3519874>
- [25] R. Morales, B. Iglar, S. Böhm, and P. Chitchaipoka, "Context-aware mobile language learning," *Procedia Computer Science*, vol. 56, pp. 82–87, 2015. Available: <https://doi.org/10.1016/j.procs.2015.07.198>
- [26] A. Developers, "Create and monitor geofences," Android Developers. Available: https://developer.android.com/develop/sensors-and-location/location/geofencing?spm=a2ty_o01.29997173.0.0.c425c921VOeNMY. Accessed on Feb. 02, 2026.
- [27] C. Groza, A. Dumitru-Cristian, M. Marcu, and R. Bogdan, "A Developer-Oriented Framework for Assessing Power Consumption in Mobile Applications: Android Energy Smells Case Study," *Sensors*, vol. 24, no. 19, article 6469, 2024. Available: <https://doi.org/10.3390/s24196469>
- [28] S. Pasricha, R. Ayoub, M. Kishinevsky, S. K. Mandal, and U. Y. Ogras, "A survey on energy management for mobile and IoT devices," *IEEE Design & Test*, vol. 37, no. 5, pp. 7–24, 2020. Available: <https://doi.org/10.1109/MDAT.2020.2976669>
- [29] A. Pathak, Y. C. Hu, and M. Zhang, "Where is the energy spent inside my app?: Fine grained energy accounting on smartphones with Eprof," in *Proceedings of the 7th ACM European Conference on Computer Systems*, pp. 29–42, 2012. Available: <https://doi.org/10.1145/2168836.2168841>
- [30] Y. Bai, H. Wang, and S. Gao, "Design Scheme for Power Consumption Optimization of Mobile Applications," in *Proceedings of the 2024 5th International Conference on Urban Construction and Management Engineering (ICUCME 2024)*, pp. 455–461, 2024. Available: https://doi.org/10.2991/978-94-6463-516-4_48
- [31] M. A. Hoque, M. Siekkinen, K. N. Khan, Y. Xiao, and S. Tarkoma, "Modeling, profiling, and debugging the energy consumption of mobile devices," *ACM Computing Surveys*, vol. 48, no. 3, article 39, 2016. Available: <https://doi.org/10.1145/2840723>
- [32] Y. Wang et al., "A framework of energy efficient mobile sensing for automatic user state recognition," in *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services*, pp. 179–192, 2009. Available: <https://doi.org/10.1145/1555816.1555835>
- [33] C. A. C. Ojeda, "Levene's Test for Verifying Homoscedasticity Between Groups in Quasi-Experiments in Social Sciences," *South Eastern European Journal of Public Health*, vol. XXV S1 2024, pp. 2119–2125, 2024. Available: <https://doi.org/10.70135/seejph.vi.2342>
- [34] J. L. Gastwirth, Y. R. Gel, and W. Miao, "The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice," *Stat. Sci.*, vol. 24, no. 3, pp. 343–360, 2009. Available: <https://doi.org/10.1214/09-STS301>
- [35] J. Brooke, "SUS: A 'Quick and Dirty' Usability Scale," in *Usability Evaluation in Industry*, 1996, pp. 189–194. Available: <https://doi.org/10.1201/9781498710411-35>

APPENDIX

A. Ethical Compliance Statement

All participants provided written informed consent. No personally identifiable information was stored; raw logs were anonymized and destroyed after computing group aggregates. Data minimization principles were enforced: only battery percentage changes, motion states, and interaction counts were collected. Participants could withdraw at any time.

B. Participant Compliance and Demographics

Table B1 reports the longitudinal compliance of the original cohort. An ‘active user’ required ≥ 1 battery log entry during 08:00–20:00. While three participants in the Control Group became non-compliant by Day 14, they were replaced by individuals from the pre-recruited buffer cohort who completed the identical 14-day protocol. To maintain maximum statistical power and ensure a balanced comparison, the results in Table B2 are derived from the full per-protocol dataset ($N = 60$), incorporating the complete data logs of these replacement users to ensure a consistent sample size ($n=30$ per group) across all metrics.

Table B1. Daily compliance rates by group

Study Day	Control (Active/Total)	Experimental (Active/Total)
Day 1	30/30 (100.0%)	30/30 (100.0%)
Day 7	28/30 (93.3%)	30/30 (100.0%)
Day 14	27/30 (90.0%)	30/30 (100.0%)

Note: The (Active/Total) ratio reflects the daily status of the initial participants. In instances of non-compliance (Control, Day 7 and 14), buffer participants were utilized to ensure the final statistical analyses (Table B2) remained based on a complete $N=60$ dataset.

Table B2 confirms baseline equivalence. No significant differences existed at pre-test (all $p > 0.05$).

Table B2. Baseline demographics (group aggregates)

Characteristic	Experimental (M $\pm$ SD)	Control (M $\pm$ SD)
Age (years)	19.1 $\pm$ 0.7	19.0 $\pm$ 0.8
Pre-Test Score (%)	42.1 $\pm$ 2.9	41.8 $\pm$ 3.2
Gender (Male Proportion)	0.47 $\pm$ 0.50	0.47 $\pm$ 0.50

C. Full Statistical Results

Table C1 reports the primary performance outcomes for the 14-day study period. Statistical significance was determined using independent-samples t-tests ($df = 58$) with pooled variance, assuming equal variances as confirmed by Levene’s Test ($p > 0.05$). All calculations were performed on the final per-protocol dataset ($N = 60$), comparing the mean daily battery drain, cumulative notification acceptance rate, and post-study SUS usability scores between the Experimental (SCT) and Control (Standard Geofence) groups.

Table C1. Group comparison of system performance metrics ($N = 60$; $df = 58$)

Metric	Experimental (M $\pm$ SD)	Control (M $\pm$ SD)	Difference	$t(58)$	p
Battery Drain (%/day)	11.52 $\pm$ 1.8	19.30 $\pm$ 2.1	-7.78	15.41	<0.001
Acceptance Rate (%)	89.44 $\pm$ 4.2	32.48 $\pm$ 6.7	+56.96	39.45	<0.001
SUS Score	81.25 $\pm$ 6.8	62.42 $\pm$ 12.1	+18.83	7.43	<0.001

D. System Usability Scale (SUS) Instrument

The following 10-item Likert scale was administered to all participants in both groups at the conclusion of the 14-day study. Responses were recorded on a scale from 1 (Strongly Disagree) to 5 (Strongly Agree) as shown in Table D1.

Table D1. SUS questions and valence

No.	Question Item	Valence
1	I think that I would like to use this system frequently.	Positive
2	I found the system unnecessarily complex.	Negative
3	I thought the system was easy to use.	Positive
4	I think I would need the support of a technical person to be able to use this system.	Negative
5	I found the various functions in this system were well integrated.	Positive
6	I thought there was too much inconsistency in this system.	Negative
7	I would imagine that most people would learn to use this system very quickly.	Positive
8	I found the system very cumbersome to use.	Negative
9	I felt very confident using the system.	Positive
10	I needed to learn a lot of things before I could get going with this system.	Negative