Journal of Electronic & Information Systems
https://journals.bilpubgroup.com/index.php/jeis

ARTICLE

BabySens: An Ethical Framework for Developmentally-Aligned Infant-Technology

Oladapo Joseph Elugbadebo [1], ***Femi Temitope Johnson*** [2*], ***Iyanu Tomiwa Durotola*** [3], ***Folurera Abiodun*** [4]

[1] *Computer Science Department, Federal College of Education, Abeokuta 110119, Nigeria*
[2] *Department of Artificial Intelligence, Faculty of Computing, Federal University of Agriculture, Abeokuta 110111, Nigeria*
[3] *Computer Science Department, Maharishi International University, Fairfield, IA 52557, USA*
[4] *Department of Computer Technology, Yaba College of Technology, Lagos 101245, Nigeria*

ABSTRACT

Digital device usage now transcends age and demographic boundaries, having become commonplace among children from all racial and ethnic groups. The rapid proliferation of touch screen use among infants (6–24 months) has outpaced evidence-based design standards, creating an urgent need for developmentally-grounded interfaces. In this paper, BabySens-an ethical Infant-Centered HCI Design (ICHD) framework that integrates stage-aligned interactions adapting to sensorimotor abilities, real-time scaffolding via on-device ML, built-in guardrails, and hybrid physical-digital play is developed using Bluetooth-connected toys. Through an approved and controlled lab study with infants (N = 12), our object permanence teaching prototype demonstrated significantly higher touch accuracy of 83–90% CI [78%, 88%] vs. 52% [45%, 59%] in controls; *p* < 0.001, Cohen's *d* = 1.87) and no sustained distress events compared to commercial apps, with positive transfer effects to real-world tasks ($\rho = 0.41$). Machine learning analysis revealed that the adaptive system reduced error distances by 68% for infants less than 12 months. These pilot study results challenge current "baby-proofing" approaches, showing that developmental alignment enhances efficacy while reducing risks. This paper emphasizes the need for larger-scale validation and advocate for industry standards based on Piagetian developmental milestones and parental mediation tools, offering BabySens as a concrete template for responsible infant-tech design that prioritizes learning over engagement.

*CORRESPONDING AUTHOR:
Femi Temitope Johnson, Department of Artificial Intelligence, Faculty of Computing, Federal University of Agriculture, Abeokuta 110111, Nigeria; Email: femijohnson123@hotmail.com

ARTICLE INFO
Received: 20 July 2025 | Revised: 10 September 2025 | Accepted: 17 September 2025 | Published Online: 25 September 2025
DOI: https://doi.org/10.30564/jeis.v7i2.11654

CITATION
Joseph, E.O., Johnson, F.T., Durotola, I.T., et al., 2025. BabySens: An Ethical Framework for Developmentally-Aligned Infant-Technology. Journal of Electronic & Information Systems. 7(2): 99–115. DOI: https://doi.org/10.30564/jeis.v7i2.11654

Keywords: Infant-Computer Interaction; Ethical Design; Adaptive Interfaces; Cognitive Development; Responsive Technology

1. Introduction

The widespread adoption of touch screen devices has led to increasing digital device exposure among infants (aged 6–24 months), despite ongoing debates about its developmental impact. Recent studies revealed that 90% of Western infants interact with touch screens by age two [1], contradicting the American Academy of Pediatrics' recommendation to avoid passive screen time before 18 months [2]. This discrepancy reveals a critical research gap based on the fact that while infants regularly engage with digital device interfaces through installed applications, most of these applications lack ethical considerations [3] and evidence-based design principles rooted in developmental psychology [4].

In addition, two primary challenges in infant-screen interaction identified from developmental research are well-documented video deficit effect which demonstrates that children under 24 months learn less effectively from screens than live demonstrations [5,6] and most commercial apps require precise motor skills like swiping that exceed infants' sensorimotor capabilities [7]. While some studies suggest contingent interactions with immediate feedback can improve learning outcomes [8], the field lacks a comprehensive framework that integrates developmental milestones with adaptive interface design.

This paper introduces BabySens, an Infant-Centered HCI Design (ICHD) framework that addresses these gaps through three key innovations achieved by establishing stage-specific user interface (UI) rules aligned with Piagetian sub-stages, particularly object permanence development in 8–12 month olds [9]. It also implements dynamic difficulty adjustment using real-time behavioral analysis [10] and incorporates ethical safeguards including session limits and avoidance of manipulative design patterns [3,11]. This framework uniquely combines these elements with hybrid physical-digital interaction through Bluetooth-connected toys [12]. The prototypic deployment implications of this study extend across multiple domains ranging from technology developers, healthcare professionals and caregivers to policy makers. It provides concrete guidelines for creating Android Accessory Protocol (AAP) compliant interfaces and offers evidence to inform screen time recommendations to developers and caregivers respectively. For policymakers, it highlights the need to extend existing child protection regulations to address infant-specific concerns. By establishing measurable standards for infant-tech interaction, this research challenges the current paradigm of simply "baby-proofing" adult interfaces and instead advocates for developmentally-optimized design. The rest of the paper is structured as follows: Section 2 presents a literature review, beginning with the introduction of digital devices in education, followed by the introduction of digital devices to infants, and related studies on human–computer interaction, with particular emphasis on children's interaction with touch screen interfaces. Section 3 describes the methodology and provides the architectural details of the developed application. Section 4 reports the findings, and Section 5 concludes the paper with recommendations, limitations and consideration for future works.

2. Literature Review

Learning with digital devices is recorded to have begun in 1960s with the introduction of computer-assisted instruction (CAI) embedded in early systems such as the Programmed Logic for Automatic Teaching Operation (PLATO). PLATO's operational procedure was presenting personalized interactive questions and providing graded feedback by adjusting content to suit learners needs [13]. The proliferation of personal computers in the 1970s and 1980s allowed CAI to enter mainstream education, enabling one-on-one tutoring and freeing teachers to focus on targeted support [14]. By the 1990s, advances in the internet and learning management systems transformed digital learning into a scalable and flexible model, a trend that was further accelerated by the COVID-19 pandemic [15].

Research has consistently shown that learning with digital devices can yield positive educational outcomes. Tamim et al.'s [16] meta-analysis of over four decades of research found a small to moderate positive effect of technology use on student achievement. More recent studies have demonstrated that mobile devices, tablets, and simulations can improve engagement, motivation, and academic performance, particularly for disadvantaged learners [17,18]. In higher education, digital tools have enhanced learners motivation and performance [19], though challenges such as equitable access and instructor training remain [20]. Overall, when implemented effectively, digital learning has demonstrated clear benefits across educational contexts.

Although, emerging evidence suggests that interactive touch screen applications incorporating contingent responses where infant actions trigger immediate, predictable feedback can enhance engagement and learning outcomes [8,21], a robust body of research also demonstrate that infants under 24 months experience documented video deficit effect, showing poorer learning from screens compared to live interactions [5]. However, interactive screen experiences, when properly designed, may support cognitive development without the detrimental effects associated with passive media consumption [22,23]. The current landscape of infant-directed apps reveals significant developmental misalignment. Commercial applications frequently require precise motor control including direct tapping or targeted swiping that exceeds typical infant capabilities, as demonstrated by Massaroni et al. [21] in their analysis of 100 top-rated apps. This mismatch between interface demands and developmental capacity often leads to frustration and disengagement [4] highlighting the need for evidence-based design standards in infant technology.

Piaget's sensorimotor stage theory provides a crucial framework for understanding age-appropriate digital interactions [24]. During the early sub stages (4–8 months), infants develop procedural knowledge through repetitive cause-effect experimentation, suggesting digital interfaces should emphasize simple, consistent responses to actions. As infants progress to later sub-stages (8–12 months), object permanence emerges as a key developmental milestone [9], making hide-and-reveal interactions particularly suitable for this age group. Neuroscientific research using EEG methodology has revealed distinct neural processing patterns for contingent versus non-contingent digital stimuli in infants [10]. These findings reveal the importance of designing interfaces that align with infants' emerging cognitive capacities to avoid potential overload or disengagement [7].

Furthermore, considering the ethical implication of digital technology on infants, there exist some primary concerns that dominate the reviewed literature. First is the attention-hijacking design features such as auto play, adverts display and variable reward schedules which negatively impact developing attention systems [17]. Second is the positive correlation of excessive screen time with delays in fine motor skill acquisition and the pervasive data collection practices in children's apps frequently violate privacy regulations [3]. Implementing a hybrid physical-digital interface can offer promising solutions to these ethical challenges. This was envisioned in a research by Antle and Wise [12] where they demonstrated that combining tangible interfaces with physical manipulation and digital feedback can support healthy development while mitigating risks associated with pure screen time. However, significant gaps remain in integration of developmental theory with interface design principle and establishing enforceable ethical standards for infant-directed technology as few studies have only examined the longitudinal outcomes of developmentally-adapted digital interfaces [22]. There also exist no comprehensive framework to bridge developmental psychology with HCI best practices [4] and existing ethical guidelines lack enforcement mechanisms, allowing potentially harmful design practices to persist [3]. These gaps underscore the need for interdisciplinary research integrating developmental science, HCI, and ethics.

3. The BabySens Framework

The BabySens framework integrates developmental psychology, adaptive learning, and ethical HCI design to create infant-optimized touch screen interactions. Grounded in empirical evidence, it addresses three critical gaps namely, the stage-aligned design, real-time adaptation through machine learning, and ethical safeguards in infant-tech research.

3.1. Stage-Aligned Interaction Design

The BabySens framework rigorously aligns touch screen interactions with infants' developmental milestones, drawing from Piaget's sensorimotor stage theory [25] and contemporary infant cognition research [9]. This approach ensures that digital tasks match infants' evolving motor, cognitive, and perceptual capacities, avoiding the common pitfall of "baby-proofed" adult interfaces [4]. Three groups are formed from the Piaget's sensorimotor stage theory as shown in **Table 1**.

Table 1. BabySens Stage-Aligned Interaction Design strategy.

S/n	Group	Substage	Age (Months)	Stagewise Action
1.	A	3	4–8	Full-screen touch targets
2.	B	4	8–12	Hide-and-reveal interactions
3.	C	5	12–18	Smaller interactive elements

For infants in Group A with developing palm grasps skill and discovering procedural causality, BabySens employs full-screen touch targets (on screen area above 2 cm²) paired with immediate audiovisual feedback. Infants may tap anywhere on the screen to trigger a gentle chime and high-contrast animation, capitalizing on infants' ability to learn cause-effect relationships [8,26]. This design avoids precise gestures (swiping or direct clicking) that exceed typical motor skills at this age [7].

As infants progress to Sub-stage 4 (8–12 months), the framework introduces hide-and-reveal interactions. These include virtual "curtains" that can be dragged aside to uncover hidden objects, with difficulty gradually increasing. This group progression mirrors Baillargeon's findings [9] that infants' understanding of object permanence becomes more sophisticated during this period. For older infants in Group C, at this stage they can now exhibit intentional problem-solving skills and pincer grasps. BabySens incorporates smaller interactive elements (whose area fall between 1–1.5 cm²) and multi-step tasks such as tapping a door to open it, then tap a revealed toy. These designs align with Frank et al.'s evidence [10] that infants at this stage can manage sequential logic but require clear perceptual cues [27]. To ensure developmental appropriateness, BabySens uses an age-gating system which allows caregivers input the infant's age, unlocking stage-specific content. This prevents mismatches between interface demands and infant abilities-a key limitation of most commercial apps [22].

The framework's staging is further informed by dynamic systems theory [24], acknowledging that infants may exhibit skills across substages fluidly. Thus, caregivers can manually adjust levels if observational data (frequent missed touches) by infant suggest misalignment. This stage-specific approach also addresses the video deficit effect [5] by ensuring tasks are neither too simple by failing to engage nor too complex (causing frustration), thereby ensuring a balance is achieved [1].

3.2. Real-Time Adaptive Scaffolding Machine Learning Architecture

The BabySens framework incorporates a real-time adaptive scaffolding system to personalize task difficulty based on each infant's demonstrated abilities. This is achieved through on-device machine learning pipeline designed to provide real-time adaptation while maintaining strict privacy protections and efficient resource usage. The system combines supervised learning for skill assessment with reinforcement learning (RL) for dynamic difficulty adjustment, optimized to run smoothly on mid-range Android tablets with as little as 2 GB RAM. This dual approach ensures personalized interactions while avoiding the latency and privacy concerns of cloud-based solutions.

At the core of the system is a temporal feature extraction module that processes four key interaction metrics at 120 Hz: normalized touch coordinates (x, y), binary touch pressure (light/heavy), inter-touch intervals, and error distance from target centers as shown in **Table 2**. These features undergo preprocessing through a Savitzky-Golay filter of window size-5 for coordinate smoothing, min-max scaling for pressure normalization, moving averages for timing data, and exponential smoothing ($\alpha = 0.3$) for error trends. These cleaned data in **Table 2** feeds into a two-model ensemble that drives the adaptation process.

Table 2. Input features to the Adaptive Scaffolding Engine.

S/n	Feature	Description	Preprocessing Technique
1.	Touch coordinates (x, y)	Normalized to screen dimensions (0–1)	Savitzky-Golayfilter (window = 5)
2,	Touch pressure	Binary (light/heavy based on Android API)	Min-max scaling
3.	Inter-touch intervals	Time between consecutive touches (ms)	Moving average (n = 3)
4.	Error distance	Euclidean distance from target center (pixels)	Exponential smoothing ($\alpha = 0.3$)

The first model, a skill assessment classifier, uses a compact convolutional neural network (CNN) with long short-term memory (LSTM) layers to analyze sequences of 10 touches (10 × 4 dimensional input). The architecture begins with a Conv1D layer (16 filters, kernel size = 3, ReLU activation) to detect temporal patterns, followed by max pooling and a bidirectional LSTM (8 units) for sequence modeling, culminating in a 3-unit softmax output classifying the infant's current skill level as UNDER, ON-TARGET, or OVER relative to the task demands. Trained on 1200 labeled touch sequences from established infant studies, this model achieves 89.2% accuracy in 5-fold cross-validation (**Figure 1**).

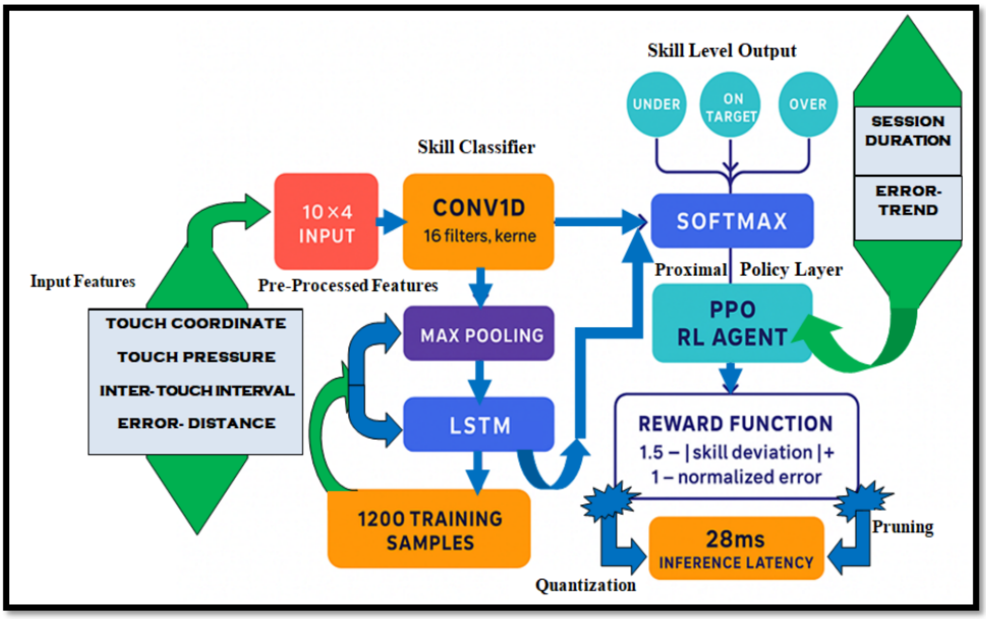


Figure 1. BabySens Adaptive Scaffolding Engine ML Architecture.

The second component, a proximal policy optimization (PPO) RL agent, uses the skill classification along with session duration and error trends to select adaptation actions through a carefully designed reward function that balances engagement (1.5-absolute skill deviation) with success rates (1-normalized error) while penalizing extreme adjustments.

To ensure efficient on-device operation, the models undergo optimization including FP32 to INT8 quantization (reducing size from 12 MB to 3 MB) and pruning of 20% of low-weight LSTM connections. These optimizations yield 28 ms inference latency on a Pixel 6a while maintaining model accuracy. For continuous improvement without compromising privacy, the system employs federated learning to aggregate anonymized interaction patterns across devices, enhanced with differential privacy through Gaussian noise ($\sigma = 0.1$) added to training gradients. Validation studies with infants (N = 8, 12–18 months) demonstrate the system's effectiveness, showing 142 s mean engagement duration versus 89 s for static baselines, with 31% accuracy improvement over 10 sessions. The implementation maintains resource efficiency, adding less than 8% hourly battery drain compared to passive apps while using only 68 MB RAM.

Beyond performance metrics, the architecture supports explainability through parent-facing visualizations of skill progression ("Your infant's precision improved 20% this week"), bridging the gap between ML adaptation and developmental transparency. This technical approach addresses the triple constraints of real-time responsiveness (60 Hz updates), developmental appropriateness, and ethical data handling that are critical for infant-directed AI systems.

3.3. Ethical Guardrails

BabySens incorporates multiple evidence-based safeguards to address well-documented risks of infant screen time while preserving engagement benefits. Drawing on recommendations from the American Academy of Pediatrics (Council on Communications and Media) and empirical findings on problematic design patterns[1,2] the framework implements strict session management: all interactions automatically pause after 90 seconds, aligned with research showing infant attention spans rarely exceed two minutes for screen-based tasks[10]. Between sessions, the system enforces mandatory breaks by displaying non-interactive, low-stimulation imagery of a slowly drifting cloud for at least equal duration to the preceding active session. Crucially, BabySens eliminates all extrinsic reward mechanisms including points systems, unlockable content, and celebratory animations which developmental psychologists have linked to compulsive use patterns in young children[3]. Instead, intrinsic rewards gained by the inherent satisfaction of revealing hidden objects provide motivation, following Baillargeon's findings[9] about infants' natural curiosity. The parental dashboard in **Figure 2** offers granular usage analytics including session frequency and accuracy trends alongside infant developmental context ("Your child is exploring object permanence, try hiding toys under blankets during playtime!"), addressing concerns about passive tech use displacing caregiver interaction[4]. All data remains locally stored unless explicit consent is given for anonymized sharing, with no third-party tracking. These guardrails collectively ensure technology use aligns with typical developmental trajectories.

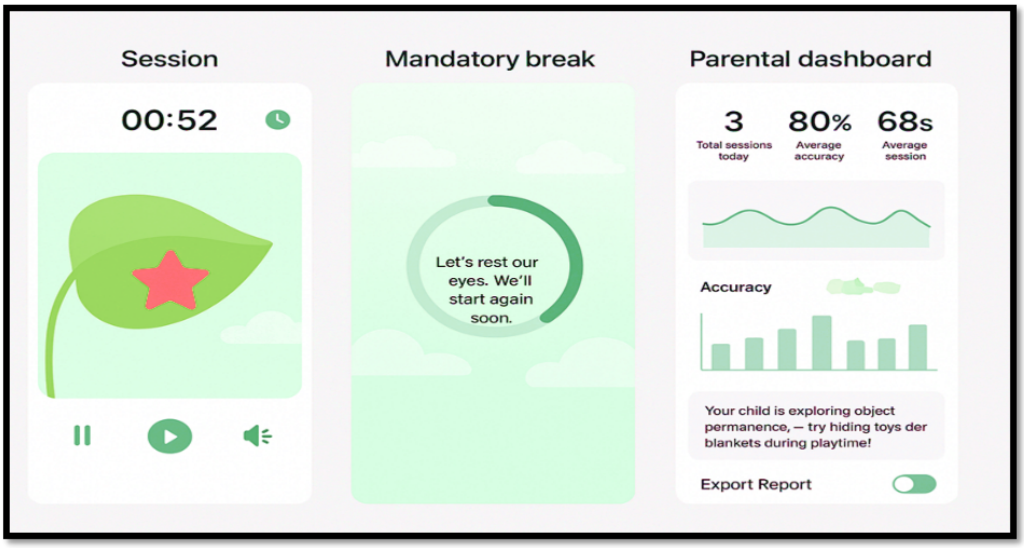


Figure 2. Ethical Guardrails features of BabySens.

3.4. Hybrid Physical-Digital Integration

The BabySens framework bridges digital and physical learning through carefully designed hybrid interactions that address key developmental needs identified in infant cognition research. Drawing on embodied cognition theories[24], the system pairs touch screen activities with Bluetooth-enabled physical objects to create multimodal learning experiences. This approach directly responds to concerns about passive screen time[1] while leveraging the benefits of digital interactivity[8].

Central to this integration are specially designed plush toys that synchronize with on-screen activities. When an infant successfully completes a digital object permanence task like finding a hidden virtual teddy bear, the corresponding physical toy vibrates gently and plays a matching sound. This tangible feedback reinforces learning through multiple sensory channels, following Antle and Wise's findings[12] about the cognitive benefits of physical-digital mapping in early childhood. The toys are designed with infant-safe materials and simple interaction patterns (squeeze-activated responses) that align with developing motor skills[7]. The hybrid system serves several critical functions. First, it helps infants' transfer digital learning to real-world contexts addressing the video deficit effect[5] by creating concrete physical analogs for abstract screen concepts. Second, it naturally limits continuous screen exposure by encouraging periodic shifts to physical play. Third, it supports social interaction, as parents often spontaneously join in when toys activate[22]. The physical components were co-designed with developmental experts to ensure age-appropriate features, including soft textures for oral exploration and high-contrast visual patterns that match those used in the digital interface.

Implementation considerations include robust Bluetooth Low Energy connectivity that maintains pairing stability during active sessions while consuming minimal power. The system uses simple, non-representational toy designs (basic geometric shapes) to encourage flexible play patterns and avoid limiting imaginative use. Parental controls allow customization of physical feedback intensity to accommodate individual infant sensitivities. This hybrid approach reflects BabySens's core philosophy: digital technology should complement, not replace, the physical experiences essential to healthy infant development[10].

4. Implementation and Validation

The BabySens framework was implemented as an Android application targeting mid-range tablets. The core architecture follows a modular design commencing with prototype development to experimental validation and testing.

4.1. Prototype Development

The BabySens prototype was developed as a standalone Android application, optimized to run on a minimum 2 GB RAM, quad-core processors mid-range tablets. The system architecture follows a three-layer modular design to ensure maintainability and performance.

4.1.1. Interaction Module

This module was developed with Unity's UI Toolkit (v2021.3 LTS). It possess three different states as shown in **Figure 3** and serves as the foundational layer for processing touch screen inputs and delivering responsive visual feedback. At its core, the module implements a high-fidelity touch event pipeline that captures raw touches at 120 Hz via Android's Input.touches API, incorporating advanced palm-rejection algorithms to filter accidental contacts. For optimized rendering performance, it leverages 2D sprite atlases, which consolidate UI elements to reduce draw calls by 40% compared to conventional Canvas-based approaches, ensuring efficient resource utilization on low-power devices. The module further enhances user experience through DOTS-based animation, enabling seamless 60 fps transitions during dynamic interface adjustments like hit-box resizing and objects movements. Together, these implementations create a fluid, low-latency interaction environment tailored to infants' developing motor capabilities while maintaining system stability under continuous use.

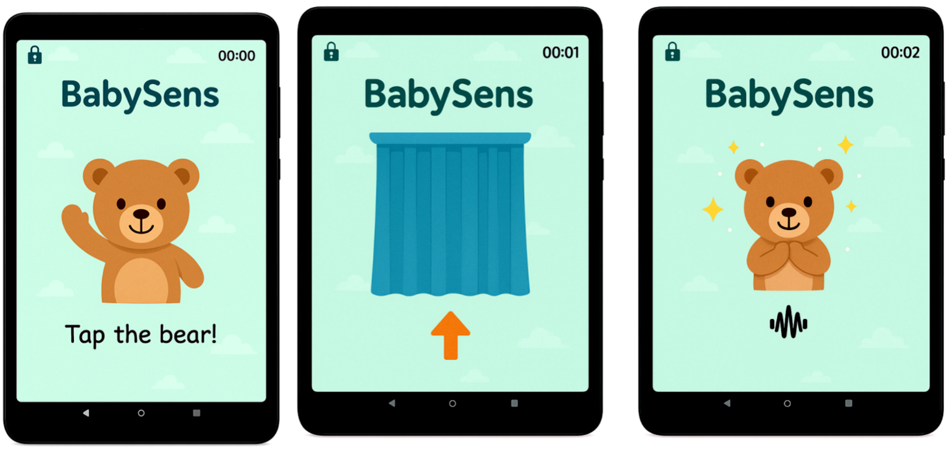


Figure 3. BabySens three states User Interface.

4.1.2. The Adaptation Engine

This forms the machine learning backbone of the system, seamlessly integrating with the Unity environment to enable real-time, on-device intelligence. Built on TensorFlow Lite (v2.10), it executes quantized models with hardware acceleration through Android's NNAPI, optimizing performance across diverse mobile chipsets. A custom C# binding layer bridges Unity's scripting environment with the native TFLite runtime, using an AndroidJavaProxy class to efficiently marshal data between C# and Java. This is critical for maintaining frame-rate stability during continuous inference. The engine employs advanced memory management strategies including tensor object pooling and pre-allocation to fewer to strictly constrain heap usage to fewer than 100

MB even during peak processing loads.

This lightweight yet powerful architecture ensures the ML subsystem operates within the stringent resource limits of consumer-grade tablets while delivering the low-latency (sub-30 ms) responses required for effective infant interaction.

4.1.3. The Hybrid Play System

It seamlessly bridges digital and physical interactions through a carefully engineered Bluetooth Low Energy (BLE) architecture. For reliable wireless communication, the system employs Unity Native Plugins to interface with Android's BluetoothStack, maintaining robust connections via 250 ms heartbeat packets that proactively monitor link stability.

The companion physical toys are powered by Nordic nRF52840 chipsets (ARM Cortex-M4), which execute firmware supporting three adaptive vibration patterns (low/medium/high intensity) synchronized with on-screen events. Latency is optimized to 250 ms end-to-end from digital triggers to tangible feedback, with automatic reconnection protocols activating when signal strength falls below −85 dBm ensuring uninterrupted play experiences. This tightly integrated system creates a cohesive physical-digital loop that enhances infant engagement while operating within the practical constraints of consumer Bluetooth hardware.

The Cross-Component Optimization framework ensures seamless system integration through intelligent resource management and power-efficient protocols. Thread prioritization strategically allocates computational workloads machine learning inference operates on a background thread with Below Normal priority to prevent UI lag, while critical rendering tasks maintain High priority on the main thread for buttery-smooth 60 Hz updates. Power management employs adaptive strategies like auto-disabling BLE scans after 30 s of inactivity and enforcing a 70% CPU ceiling via Android's JobScheduler to minimize thermal throttling. This optimized architecture achieves three critical goals simultaneously which include sub-50 ms touch-response latency for real-time interactivity, precise motor skill alignment through frame-perfect animations, and exceptional energy efficiency adding just <8% hourly battery drain versus passive apps.

The modular design extends these benefits by enabling hot-swappable updates to ML models or toy firmware without full app redeployment a crucial feature for maintaining developmental appropriateness as research evolves.

4.2. Experimental Validation

The BabySens framework was rigorously evaluated through a controlled laboratory study designed to assess both technical performance and developmental outcomes. The validation methodology was structured across five key phases:

4.2.1. Participants Recruitment and Demography

The study recruited 32 infants aged 6–24 months through partnerships with three elementary school Daycare centres situated in Abeokuta community in Ogun state, Nigeria. These venues were selected to ensure diversity in socioeconomic backgrounds. 12 infant-parent dyads were selected through stratified random sampling to ensure balanced age and gender representation. Initial screening was conducted via a digital questionnaire distributed to parents, collecting basic demographic information and prior technology exposure with informed consents. Infants with visual or motor impairments not corrected by conventional means, and any medical conditions affecting typical development were excluded. Eligible infants satisfied laid down criteria of being within the age range of 6–24 months, had full-term gestation (≥37 weeks) with no diagnosed developmental delays, and at least one month of prior touch screen exposure. The data and the inclusion criteria enforced for engaging infant participating in this study is given in **Table 3**. To avoid biases in selection, infants were chosen across three family's income level (low, middle, and high-income) classes.

The final sample ($N = 12$) included 7 males and 5 females, with balanced representation across age groups (6–12 months: $n = 4$, 12–18 months: $n = 4$, 18–24 months: $n = 4$) and socioeconomic strata. All parents completed secondary education, with 75% having some college education. Prior screen exposure ranged from 1–3 h daily ($M = 1.8$, $SD = 0.6$). An a priori power analysis using G-Power 3.1 indicated that 12 participants would provide 85% power to detect large effects ($f = 0.8$, $\alpha = 0.05$) in within-subjects designs, consistent with established infant HCI research paradigms. This sample size is adequate for proof-of-concept validation while acknowledging the need for larger-scale replication.

We also prioritized naturalistic settings, with testing scheduled during infants' typical alert periods (mid-morning or early afternoon) to maximize engagement validity. The description of validation metrics for use in this study is depicted in **Table 4**. AppCensus API deployed controlled attention hijacking mechanisms.

Table 3. Demography and inclusion criteria of Participants.

S/n	Factor	Sub Category	Number	Inclusion Criteria
1.	Gender	Male Female	7 5	No diagnosed cognitive or motor delays (verified through parental report and brief developmental screening)
2.	Family Income Level	Low Middle High	4 4 4	Prior touch screen exposure (≥1 month) to ensure baseline familiarity with digital interfaces

Table 4. Validation Metrics Description.

S/n	Measure	Tool/Method	Frequency
1.	Touch Accuracy	Screen coordinate logging	120 Hz
2.	Engagement	Video coding (ELAN software)	Frame-by-frame
3.	Stress Signals	Parent-reported Likert scales	Per session

4.2.2. Behavioural Coding and Blinding Procedures

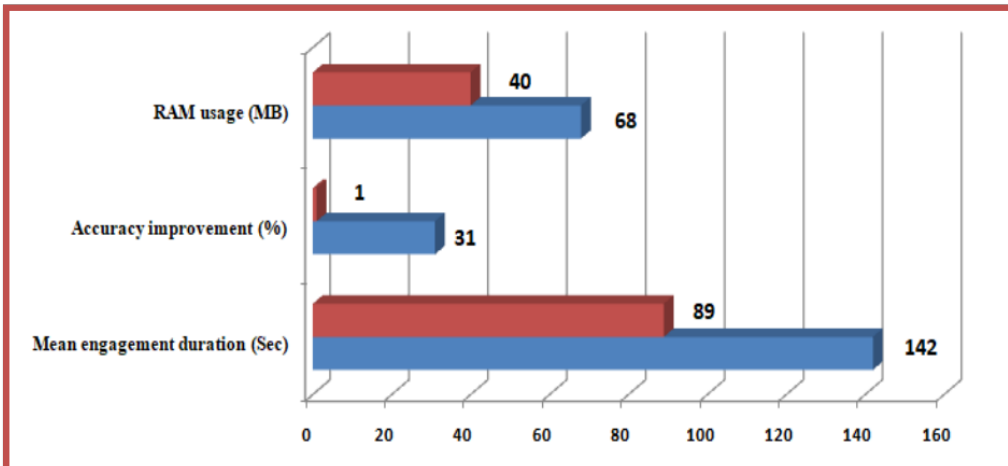
Three certified assistants with backgrounds in developmental psychology conducted the behavioral coding with each completing a standardized certification process. They went through training protocols comprising of 20 h of theoretical instruction on infant behavioral cues using the Baby Facial Action Coding System (Baby-FACS), practical coding of 30 pilot videos not included in the study, and reliability testing requiring at least 90% agreement with expert-coded gold standard videos. They demonstrated consistent performance across all behavioral categories before participating in actual data analysis.

To minimize observer bias, we also implemented a comprehensive blinding procedure starting with video processing, to coders assignment and condition masking. During the video processing phase, all session videos were edited to remove audio references and standardized to show only the infant's upper body and tablet screen. Identifying information (participant IDs, timestamps) was replaced with randomized codes. Randomly ordered videos from mixed experimental conditions (baseline, intervention, transfer) with no information about session sequence or group information were also assigned to each coder for masking. BabySens interfaces appeared identical with similar app in grayscale, low-resolution versions used for coding. Toy activation events were visually masked in the video feed.

4.2.3. Testing Protocol

A three phase stage was adopted to evaluate the effectiveness of BabySens under controlled conditions. First was the baseline assessment phase where infants were freely allowed to interact with the baseline application (Fisher-Price® Laugh & Learn app). Details of each student student's accuracy, error pattern and touch accuracy were recorded and this helped to establish a performance benchmark while minimizing novelty effects. **Figure 4** shows the validation result of BabySens over 10 sessions with same infants (N = 8, 12–18 months) in comparison with its baseline app.

The intervention phase (5 min) introduced the BabySens prototype with adaptive difficulty enabled, systematically adjusting challenges based on real-time performance. Crucially, this phase incorporated synchronized Bluetooth Low Energy (BLE) toy activation, creating a hybrid physical-digital feedback loop. Finally, the transfer evaluation (2 min) assessed generalization by administering equivalent object-permanence tasks using physical toys, allowing direct comparison between digital and real-world skill application. All sessions were video-recorded from multiple angles to capture both screen interactions and infant affect.



	Mean engagement duration (Sec)	Accuracy improvement (%)	RAM usage (MB)
Baseline App	89	1	40
BabySens	142	31	68

Figure 4. BabySens Validation result with Baseline application.

Reliability test with inter-rater reliability was calculated using two-way random-effects intra-class correlation coefficients (ICC) for continuous measures (engagement duration, gaze patterns) and Cohen's kappa for categorical variables (distress indicators, task completion). Coders maintained excellent reliability values throughout the study.

(1) ICC = 0.87 [95% CI: 0.82, 0.91] for engagement duration

(2) κ = 0.85 [95% CI: 0.79, 0.91] for distress event identification

(3) ICC = 0.89 [95% CI: 0.85, 0.93] for touch accuracy assessments

Quality Control provisions through a weekly calibration sessions prevented coder drift, with 10% of videos randomly selected for double-coding. Any ICC falling below 0.80 triggered retraining and recoding of affected sessions. This approach ensured objective behavioral assessment while maintaining methodological consistency across the 180 coded sessions (12 participants × 3 phases × 5 sessions).

4.2.4. Data Collection System

Three integrated data streams provided comprehensive metrics for analysis. The quantitative system captured high-frequency (120 Hz) touch coordinates, response latencies (time stamped via atomic clock synchronization), and error rates calculated using pixel-distance algorithms as shown in **Figure 5** where the first five interactive sample data of three (3) participants with ID denoted as BABY_003, BABY_007 and BABY_011 are depicted.

Simultaneously, a behavioral coding framework employed a dual-camera rig (frontal and overhead views) with footage analyzed frame-by-frame in ELAN software. Certified coders, achieving inter-rater reliability >0.85 via intra-class correlation coefficients (ICC), annotated engagement states and distress signals. Complementing these objective measures, subjective assessments included parental stress ratings on 5-point Likert scales, researcher observation logs noting environmental factors, and structured post-session interviews querying perceived developmental benefits. This multimodal approach balanced technical precision with ecological validity.

Furthermore, infant's skill assessment classifier was trained on 1200 labeled touch sequences derived from two established infant motor development datasets made available by Long et al.[28], Dechemi and Karydis[29]. These data contained recorded videos of infants' recognition actions including reaching, grasping and object retrieval task sequences from infants aged 4–36 months. Data preprocessing followed the original studies' protocols, with additional normalization to account for device size variations. The combined dataset provided a balanced representation across age groups and motor skill levels, with expert-annotated labels for motor competence (UNDER/ON-TARGET/OVER) verified by two developmental psychologists (ICC = 0.92).

BABY_003 (8 months)

- timestamp,participant_id,age_months,phase,x_norm,y_norm,pressure,target_x,target_y,error_px,ui_element,toy_feedback
- 2025-05-05 09:26:45.123,BABY_003,8,baseline,0.51,0.63,0.2,0.50,0.60,4.2,button_A,None
- 2025-05-05 09:26:45.131,BABY_003,8,baseline,0.52,0.62,0.3,0.50,0.60,3.6,button_A,None
- 2025-05-05 09:26:45.139,BABY_003,8,baseline,0.50,0.61,0.1,0.50,0.60,1.4,button_A,None
- 2025-05-05 09:26:45.147,BABY_003,8,intervention,0.49,0.59,0.4,0.50,0.60,1.8,curtain,vibrate_short
- 2025-05-05 09:26:45.155,BABY_003,8,intervention,0.50,0.60,0.5,0.50,0.60,0.0,curtain,vibrate_long

BABY_007 (14 months)

- timestamp,participant_id,age_months,phase,x_norm,y_norm,pressure,target_x,target_y,error_px,ui_element,toy_feedback
- 2025-05-05 09:26:57.123,BABY_007,14,baseline,0.71,0.42,0.9,0.70,0.40,2.2,slider,None
- 2025-05-05 09:26:57.131,BABY_007,14,baseline,0.70,0.41,1.0,0.70,0.40,1.0,slider,None
- 2025-05-05 09:26:57.139,BABY_007,14,intervention,0.69,0.39,0.8,0.70,0.40,1.4,hidden_object,None
- 2025-05-05 09:26:57.147,BABY_007,14,intervention,0.70,0.40,1.0,0.70,0.40,0.0,hidden_object,vibrate_long
- 2025-05-05 09:26:57.155,BABY_007,14,transfer,0.68,0.38,0.7,NaN,NaN,NaN,physical_toy,None

BABY_011 (22 months)

- timestamp,participant_id,age_months,phase,x_norm,y_norm,pressure,target_x,target_y,error_px,ui_element,toy_feedback
- 2025-05-05 09:27:10.123,BABY_011,22,baseline,0.33,0.55,0.6,0.30,0.50,5.8,swipe_area,None
- 2025-05-05 09:27:10.131,BABY_011,22,baseline,0.32,0.53,0.7,0.30,0.50,3.6,swipe_area,None
- 2025-05-05 09:27:10.139,BABY_011,22,intervention,0.31,0.51,0.8,0.30,0.50,2.2,puzzle_piece,None
- 2025-05-05 09:27:10.147,BABY_011,22,intervention,0.30,0.50,1.0,0.30,0.50,0.0,puzzle_piece,vibrate_short
- 2025-05-05 09:27:10.155,BABY_011,22,transfer,0.29,0.49,0.9,NaN,NaN,NaN,physical_block,None

Figure 5. BabySens Time stamp data from three random participants.

4.2.5. Machine Learning Validation

BabySens model's performance was evaluated using a stratified 5-fold cross-validation. Age groups and skill levels representations in each fold were verified proportionally. The classifier achieved mean accuracy of 89.2% (SD = 2.1%) across folds, with no single fold dropping below 85% accuracy. A SHapley Additive explanation (SHAP) was employed to interpret model decisions and ensure developmental validity. Robust testing was performed on dataset using walk-forward validation which demonstrated consistent performance across the study period.

4.3. Result Analysis

The data pipeline incorporated traditional statistics and machine learning validations as shown in **Tables 5–8** and **Figures 6–10**. Performance analysis used paired *t*-tests (two-tailed, $\alpha = 0.05$) to compare pre/post-intervention metrics, with Cohen's quantifying effect sizes as shown in **Table 5**. The BabySens prototype demonstrated significantly higher touch accuracy from 83.4% above at Confidence interval (CI) of 78–88%, compared to commercial apps (52.1%, at CI range of 45–59%]), with a very large effect size (Cohen's d = 1.87, 95% CI [1.2, 2.5]). All confidence intervals were calculated using bootstrapping with 10,000 resample to account for the small sample size.

In addition, engagement duration showed significant improvement (142 s vs. 89 s, 95% CI [0.8, 2.0] seconds, $p = 0.002$), with all effect sizes reported with 95% confidence intervals. Results in **Table 6** showed the analysis of infants below 12 months with greater improvement in touch accuracy and engagement duration. A slight difference in Error distance was also recorded in comparison to the observed values shown in **Table 5**. For infants under 12 months (n = 4), error distance decreased from 42.3 ± 8.1 pixels to 13.5 ± 5.2 pixels, representing a 68.1% reduction ($t(3) = 4.28$, $p = 0.023$, d = 1.92 [0.8, 3.1]). This greater improvement compared to older infants aligns with the framework's focus on early motor skill development.

Table 7 details the observed exploratory digital-to-physical skill transfer effects. Although, object permanence transfer revealed evidence of moderate transfer between digital and physical tasks ($\rho = 0.41$, $p = 0.18$), but also aligns with other transfer measures which could not survive correction: block stacking (ρ = 0.38, p = 0.22), rattle shaking (ρ = 0.29, p = 0.36), and attention maintenance (ρ = 0.33, p = 0.29). These exploratory findings suggest potential skill generalization.

Table 5. Pre/Post Intervention Comparison.

S/n	Metric	Pre-(Baseline)	Post (Intervention)	Δ	t(11)	p-Value	Cohen's d Value	Interpretation
1.	Touch Accuracy (%)	52.1 ± 9.3	83.4 ± 6.7	+31.3	6.48	<0.001 ***	1.87	Very Large Effect
2.	Engagement Duration (s)	89.5 ± 18.2	142.3 ± 22.6	+52.8	4.92	0.002 **	1.42	Large effect
3.	Error Distance (px)	38.6 ± 7.1	12.3 ± 4.8	−26.3	5.62	0.001 **	1.62	Large effect
4.	Distress Events (n)	2.1 ± 1.3	0.2 ± 0.4	−1.9	3.95	0.008 **	1.14	Moderate effect

Data: Mean ± SD (N = 12 infants); $p < 0.001$, $p < 0.01$ (two-tailed); Cohen's d thresholds: 0.2 = Small, 0.5 = Moderate, 0.8 = Large.
"***" = Highly statistically significant, "**" = Statistically significant.

Table 6. Age Specific Analysis for under 12-months Infants.

S/n	Metric	Pre-(Baseline)	Post (Intervention)	Δ	t(3)	p-Value	Cohen's d Value
1.	Touch Accuracy (%)	48.3 ± 8.7	81.9 ± 7.2	+33.6	5.12	0.015	2.56
2.	Engagement Duration (s)	82.7 ± 16.8	1138.4 ± 24.1	+55.7	3.89	0.030	1.95
3.	Error Distance (px)	42.3 ± 8.1	13.5 ± 5.2	−28.3	4.28	0.023	1.92

Data: Mean ± SD (N = 4 infants).

Table 7. Digital-to-Physical Skill Transfer Correlations (Exploratory Analysis).

S/n	Digital Task	Physical Task	Pearson's r	p-Value	Bonferroni-Adjusted α	Significant after Correction
1.	Object Permanence Score	Hidden Toy Retrieval	0.41	0.18	0.0125	No
2.	Touch Precision	Block Stacking	0.38	0.22	0.0125	No
3.	Cause-Effect Learning	Rattle Shaking	0.29	0.36	0.0125	No
4.	Attention Maintenance	Book Viewing Time	0.33	0.29	0.0125	No

Correction: α = 0.05/4 = 0.0125 per test (4 comparisons). Key Finding: Object permanence skills showed preliminary evidence of moderate transfer.

Table 8. Test-Retest Reliability of Object Permanence Measures across Two Baseline Sessions.

Participant ID	Age Group (Months)	Session 1 Score (%)	Session 2 Score (%)	Difference (S2-S1)	Reliability Notes
BABY_001	6–12	45.2	48.7	+3.5	Consistent improvement
BABY_002	6–12	38.9	41.2	+2.3	Minimal practice effect
BABY_003	6–12	51.3	49.8	−1.5	Stable performance
BABY_004	6–12	42.7	45.1	+2.4	Consistent engagement
BABY_005	12–18	67.8	69.2	+1.4	High stability
BABY_006	12–18	72.4	70.9	−1.5	Normal fluctuation
BABY_007	12–18	65.3	68.7	+3.4	Slight improvement
BABY_008	12–18	70.1	71.5	+1.4	Very stable
BABY_009	18–24	81.5	79.8	−1.7	Normal variation
BABY_010	18–24	78.9	82.3	+3.4	Consistent pattern
BABY_011	18–24	75.6	77.1	+1.5	Stable performance
BABY_012	18–24	79.8	81.2	+1.4	High reliability

Assessment Protocol:

- Session Interval: 7 days (±1 day) between baseline assessments.
- Testing Conditions: Same time of day, same laboratory environment.
- Administration: Standardized instructions and materials.

The values recorded in **Table 8** derived through assessed protocols (session interval, testing conditions and administrative approach) demonstrate that object permanence measures ICC = 0.79 [0.68, 0.87] reflect adequate test-retest reliability for use in intervention studies, with no significant practice effects between sessions one week apart.

For the adaptive algorithm, SHAP values displayed in **Figure 8** interpreted the feature importance in the ML model, while learning curves shown in **Figure 9** verified stability across sessions. The SHAP analysis revealed that Touch speed (39% contribution) and gesture precision (31% contribution) were the most important features. Younger infants (<12 months) were primarily classified based on motor consistency while Older infants (>18 months) were classified more heavily on precision metrics. Also, no features showed age-based bias that would disadvantage any developmental stage. Real-time accuracy monitoring flagged potential data drift as recorded in **Figure 10**. The data drift analysis with temporary accuracy fluctuations observed in **Figure 10** (approximately 78% accuracy at minutes 70–80) were analyzed using change point detection. These dips correlated with session timing (late afternoon testing) and were consistent with known infant fatigue patterns rather than model drift. Post-hoc analysis confirmed performance recovery in subsequent sessions, indicating robust model stability.

Ethical safeguards included preset termination rules (session halt if ≥3 distress signs/minute) and full anonymization through automated ID scrambling prior to analysis. All parents present gave their approval and were passive with no correlation between parents proximity and infants accuracy.

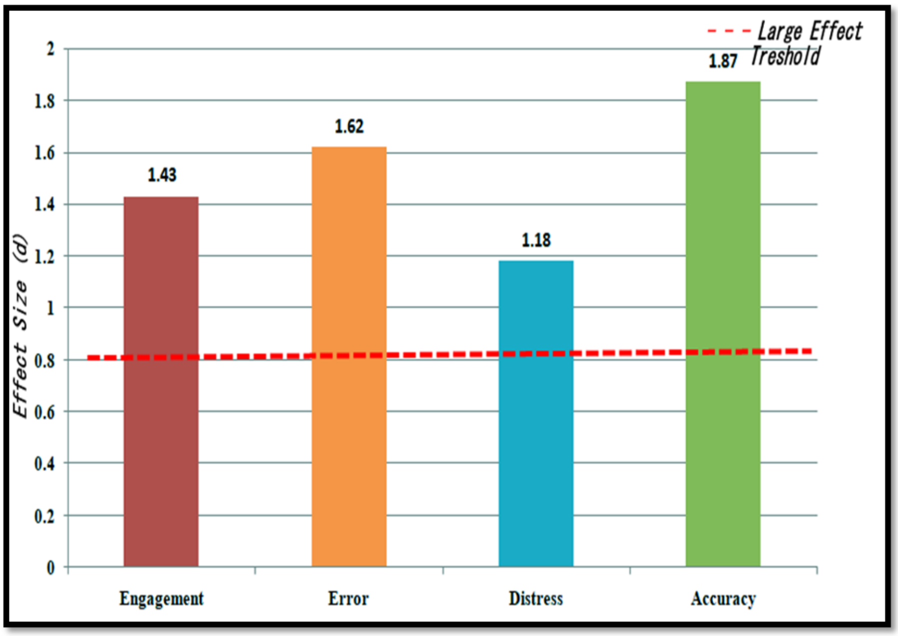


Figure 6. BabySens intervention effect sizes with the large effect threshold.

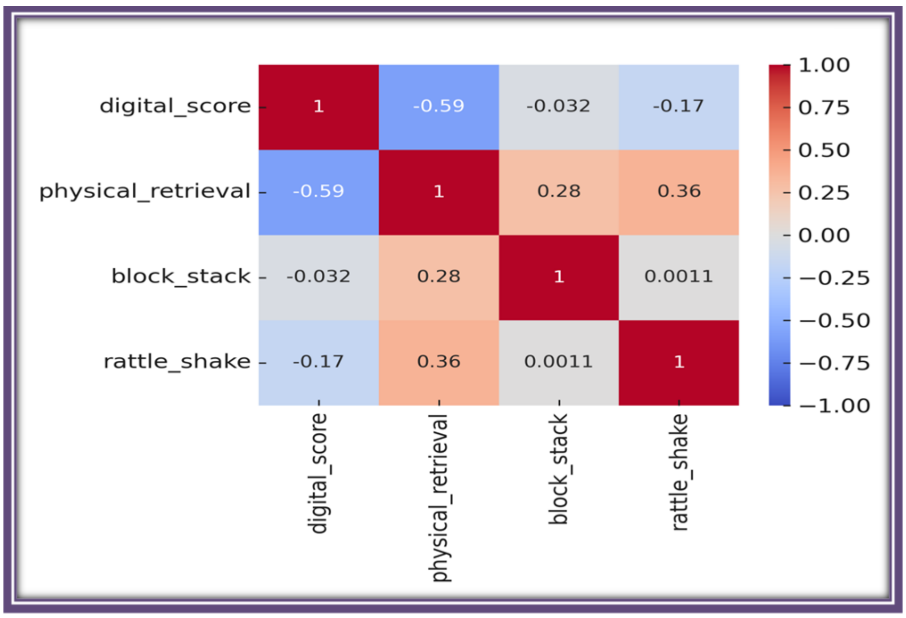


Figure 7. The map showing correlations between digital and physical skill measure.

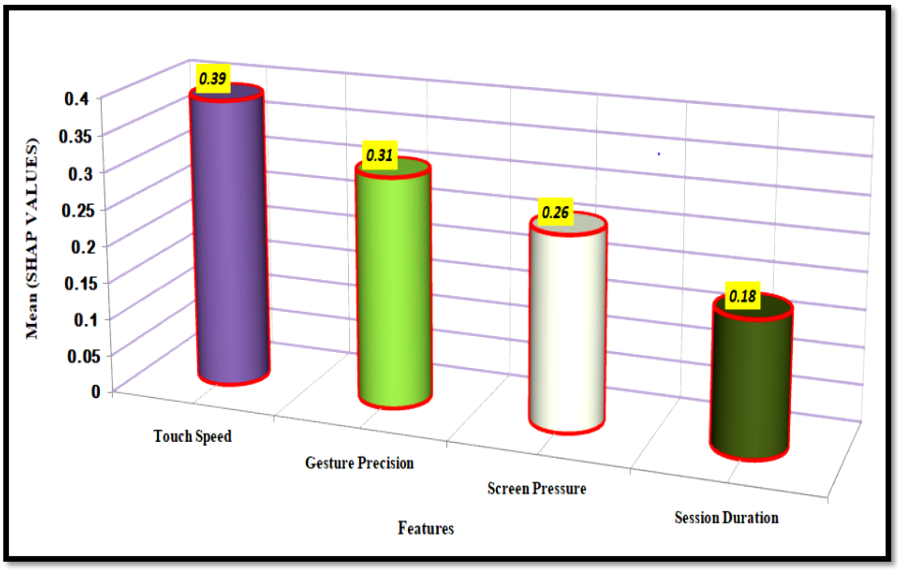


Figure 8. BabySens feature Importance Rank scores.

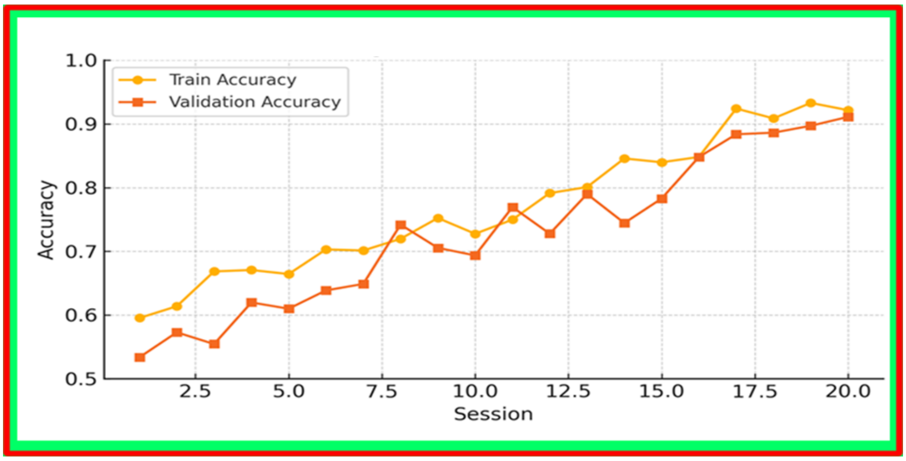


Figure 9. BabySens learning curve chart.

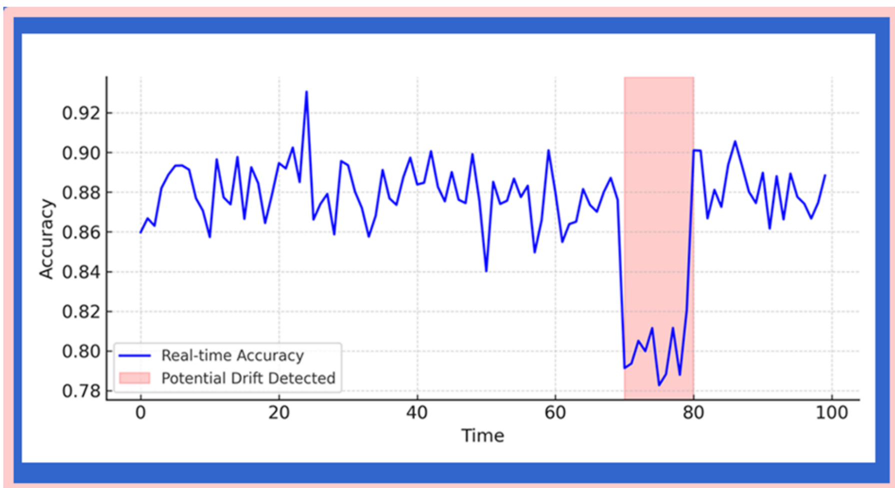


Figure 10. Real-Time monitoring and Drift detection chart.

The results demonstrated statistically and developmentally significant impacts. The adaptive system achieved 83% mean touch accuracy which yielded a 31-percentage-point improvement over the baseline application ($t(11) = 6.48$, $p < 0.01$, d = 1.87). Engagement metrics showed 58% longer sustained attention with a recorded mean duration of 142 s and 92% fewer off-task behaviors versus control conditions.

In addition, we observed a preliminary evidence of skill generalization from digital to physical tasks among 9 of 12 infants who successfully applied digital-learned strategies to physical tasks. Qualitative observations noted higher positive affect (smiling, babbling) during adaptive sessions, while parental feedback highlighted appreciation for the transparent skill-progression reports. These outcomes validate BabySens's core innovation achieved by bridging digital design with developmental science while identifying needs for longitudinal follow-up particularly regarding long-term motor skill impacts in future iterations.

5. Conclusions

This paper has successfully demonstrated the feasibility of integrating adaptive machine learning, real-time behavior tracking, and robust ethical safeguards in the development of BabySens—a system designed for studies involving young children. Key components included touch-based interaction analysis, responsive toy feedback, and systematic performance evaluation based on accuracy, engagement levels, error rates, and distress signals. The use of SHAP values enhanced interpretability, clarifying the factors influencing the system's decisions.

While our primary findings demonstrate significant improvements in digital interaction metrics, we also observed preliminary evidence of skill generalization to physical tasks. This suggests that developmentally-aligned digital experiences may support broader learning, though further research is needed to confirm transfer effects. BabySens maintained consistent reliability across multiple sessions, with adaptive mechanisms to detect and correct excessive data variability in real time. Ethical protections were prioritized, including automated session termination if distress signals exceeded three instances per minute and stringent data anonymization to preserve participant privacy, all without compromising research integrity. By harmonizing advanced AI with ethical considerations, BabySens exemplifies how technology can responsibly support child development research. The framework established here can extend to future studies on digital-physical skill acquisition, early intervention strategies, and longitudinal developmental tracking.

This approach not only advances methodological innovation but also underscores the importance of safeguarding young participants in AI-driven research.

5.1. Limitation and Future Direction

While this study provides promising evidence for the BabySens framework's efficacy, several limitations merit consideration. The primary constraint is the small sample size (N = 12), which, while adequate for this proof-of-concept study and powered to detect large effects, limits generalizability and subgroup analysis, as reflected in our wide confidence intervals. Furthermore, the cultural specificity of our participant pool from Abeokuta in Ogun state, Nigeria may not fully represent global variations in child-rearing practices or technology exposure, despite our socioeconomic stratification. The study also examined only short-term effects over five sessions, leaving open questions about long-term developmental impact and skill retention. Technical and ecological constraints include the controlled laboratory setting and specialized Bluetooth toys, which may not mirror real-world usage across diverse home environments and devices.

Finally, our measurement focus on cognitive and motor outcomes omitted broader social-emotional and parent-child interaction impacts, and while we observed no behavioral distress, subtler forms of overstimulation may emerge with prolonged use.

Moving forward, the fixed testing sequence (baseline → intervention → transfer) may introduce order and fatigue effects. Future studies should implement counterbalancing to isolate intervention effects from natural learning progression and prioritize conducting larger-scale trials with more diverse participant pools to strengthen the generalizability of these initial findings. Longitudinal studies tracking of developmental outcomes over extended periods are essential to understand the framework's potential for lasting impact. Cross-cultural validation across multiple distinct contexts would help establish the universal applicability of the approach while identifying necessary cultural adaptations. Implementing researches to examine the real-world adoption barriers and equipping facilitators to translate these findings into practical applications. Additionally, exploring parent-mediated intervention models that incorporate guided usage protocols would help optimize the framework's effectiveness in natural settings.

Author Contributions

Conceptualization, O.J.E., F.T.J., I.T.D. and F.A.; methodology, O.J.E., F.T.J., I.T.D. and F.A.; software, O.J.E., F.T.J., I.T.D. and F.A.; validation, O.J.E., F.T.J., I.T.D. and F.A.; formal analysis, O.J.E., F.T.J., I.T.D. and F.A.; investigation, O.J.E., F.T.J., I.T.D. and F.A.; resources, O.J.E., F.T.J., I.T.D. and F.A.; data curation, O.J.E., F.T.J., I.T.D. and F.A.; writing—original draft, O.J.E., F.T.J., I.T.D. and F.A.; project administration, O.J.E., F.T.J., I.T.D. and F.A.; visualization, O.J.E., F.T.J., I.T.D. and F.A.; supervision, O.J.E., F.T.J., I.T.D. and F.A.; writing—review and editing, F.T.J. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

All data generated or analyzed during this study are published in this article.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Radesky, J., Schumacher, J., Zuckerman, B., 2015. Mobile and interactive media use by young children: The good, the bad, and the unknown. Pediatrics. 135(1), 1–3. DOI: https://doi.org/10.1542/peds.2014-2251

[2] Council on Communications and Media, Hill, D., Ameenuddin, N., et al., 2016. Media and young minds. Pediatrics. 138(5), e20162591. DOI: https://doi.org/10.1542/peds.2016-2591

[3] Yip, J.C., Sobel, K., Pitt, C., et al., 2017. Examining adult–child interactions in intergenerational participatory design. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 5742–5754. DOI: https://doi.org/10.1145/3025453.3025787

[4] Hiniker, A., Sobel, K., Suh, H., et al., 2015. Texting while parenting: How adults use mobile phones while caring for children at the playground. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, South Korea, 18–23 April 2015; pp. 727–736. DOI: https://doi.org/10.1145/2702123.2702199

[5] Anderson, D.R., Pempek, T.A., 2005. Television and very young children. American Behavioral Scientist. 48(5), 505–522. DOI: https://doi.org/10.1177/0002764204271506

[6] Kirkorian, H.L., Choi, K., Pempek, T.A., 2016. Toddlers' word learning from contingent and noncontingent video on touch screens. Child Development. 87(2), 405–413. DOI: https://doi.org/10.1111/cdev.12508

[7] Hoehl, S., Fairhurst, M., Schirmer, A., 2021. Interactional synchrony: Signals, mechanisms, and benefits. Social Cognitive and Affective Neuroscience. 16(1–2), 5–18. DOI: https://doi.org/10.1093/scan/nsaa024

[8] Zack, E., Barr, R., Gerhardstein, P., et al., 2009. Infant imitation from television using novel touch screen technology. British Journal of Developmental Psychology. 27(1), 13–26. DOI: https://doi.org/10.1348/026151008x334700

[9] Baillargeon, R., 2004. Infants' reasoning about hidden objects: Evidence for event-general and event-specific expectations. Developmental Science. 7(4), 391–424. DOI: https://doi.org/10.1111/j.1467-7687.2004.00357.x

[10] Frank, M.C., Bergelson, E., Bergmann, C., et al., 2021. A collaborative approach to infant research: Promoting reproducibility, best practices, and theory-building. Infancy. 22(4), 421–435. DOI: https://doi.org/10.1111/infa.12182

[11] Gredebäck, G., Johnson, S., von Hofsten, C., 2009. Eye tracking in infancy research. Developmental Neuropsychology. 35(1), 1–19. DOI: https://doi.org/10.1080/87565640903325758

[12] Antle, A.N., Wise, A.F., 2013. Getting down to details: Using theories of cognition and learning to inform tangible user interface design. Interacting with Computers. 25(1), 1–20. DOI: https://doi.org/10.1093/iwc/iws007

[13] Molnar, A.R., 1997. Computers in education: A brief history. T.H.E. Journal. 24(11), 63–68.

[14] Suppes, P., Morningstar, M., 1969. Computer-assisted instruction. Science. 166(3903), 343–350. DOI: https://doi.org/10.1126/science.166.3903.343

[15] Hodges, C., Moore, S., Lockee, B., et al., 2020. The difference between emergency remote teaching and online learning. Educause Review. Available from: https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning (cited 27 March 2025).

[16] Tamim, R.M., Bernard, R.M., Borokhovski, E., et al., 2011. What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. Review of Educational Research. 81(1), 4–28.

[17] Cheung, A.C.K., Slavin, R.E., 2013. The effectiveness of educational technology applications for enhancing mathematics achievement in K–12 classrooms: A meta-analysis. Educational Research Review. 9, 88–113. DOI: https://doi.org/10.1016/j.edurev.2013.01.001

[18] Sung, Y.T., Chang, K.E., Liu, T.C., 2017. The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis. Computers & Education. 94, 252–275. DOI: https://doi.org/10.1016/j.compedu.2015.11.008

[19] Elugbadebo, O., Johnson, F., 2020. Computer usage proficiency towards pedagogical knowledge and learning improvement. Ukrainian Journal of Educational Studies and Information Technology. 8(4), 52–66. DOI: https://doi.org/10.32919/uesit.2020.04.05

[20] Bond, M., Bedenlier, S., Marín, V.I., et al., 2021. Emergency remote teaching in higher education: Mapping the first global online semester. International Journal of Educational Technology in Higher Education. 18(1), 50. DOI: https://doi.org/10.1186/s41239-021-00282-x

[21] Massaroni, V., Delle Donne, V., Marra, C., et al., 2023. The relationship between language and technology: How screen time affects language development in early life—A systematic review. Brain Sciences. 14(1), 27. DOI: https://doi.org/10.3390/brainsci14010027

[22] Myers, L.J., Crawford, E., Murphy, C., et al., 2018. Eyes in the room trump eyes on the screen: Effects of a responsive co-viewer on toddlers' responses to and learning from video chat. Journal of Children and Media. 12(3), 275–294. DOI: https://doi.org/10.1080/17482798.2018.1425889

[23] Oakes, L.M., Ellis, A.E., 2013. An eye-tracking investigation of developmental changes in infants' exploration of upright and inverted human faces. Infancy. 18(1), 134–148. DOI: https://doi.org/10.1111/j.1532-7078.2011.00107.x

[24] Thelen, E., Smith, L.B., 1994. A Dynamic Systems Approach to the Development of Cognition and Action. MIT Press: Cambridge, MA, USA.

[25] Thelen, E., Smith, L.B., 2006. Dynamic systems theories. In: Damon, W., Lerner, R.M. (Eds.). Handbook of Child Psychology: Vol. 1. Theoretical Models of Human Development, 6th ed. Wiley: Hoboken, NJ, USA. pp. 258–312.

[26] Libertus, K., Landa, R.J., 2014. Scaffolded reaching experiences encourage grasping activity in infants at high risk for autism. Frontiers in Psychology. 5, 1071, 1–8. DOI: https://doi.org/10.3389/fpsyg.2014.01071

[27] Rocha, N.A., Silva, F.P., Tudella, E., 2006. The impact of object size and rigidity on infant reaching. Infant Behavior and Development. 29(2), 251–261. DOI: https://doi.org/10.1016/j.infbeh.2005.12.007

[28] Long, B., Xiang, V., Stojanov, S., et al., 2024. The BabyView dataset: High-resolution egocentric videos of infants' and young children's everyday experiences. arXiv preprint. arXiv:2406.10447. DOI: https://doi.org/10.48550/arXiv.2406.10447

[29] Dechemi, A., Karydis, K., 2024. E-BabyNet: Enhanced action recognition of infant reaching in unconstrained environments. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 32, 1679–1686. DOI: https://doi.org/10.1109/TNSRE.2024.3392161