EdTech Tulna

Establishing Quality Standards in EdTech

Developed by

IIT Delhi

IIT Bombay

Personalized Adaptive Learning: Global Evidence and Policy Insights

2025

Table of Contents

Executive Summary

Since 2021, EdTech Tulna has played a pivotal role in enabling governments and educators to identify high-quality adaptive learning solutions through rigorous independent evaluations. While these evaluations have successfully elevated product standards, a recurring challenge has emerged for policymakers - *the need for evidence-informed guidance on program design*. Specifically, decision-makers require clarity on which interventions work best for specific grades and subjects, and what implementation conditions drive success.

To address this, Tulna conducted a Rapid Evidence Review of global K–12 studies. This article synthesizes insights from 73 distinct PAL studies, examining both the overall effect on learning and the critical implementation choices, such as dosage and facilitation, that influenced impact.

Key insights and their policy implications

1. **PAL interventions demonstrated consistent and comparable learning gains across both high and low-middle-income (LMIC) countries:** This validates PAL's potential to deliver improvements in developing-world contexts, with higher infrastructure and capability challenges.
2. **Across Mathematics, English, and Science, studies reported statistically significant improvements in learning outcomes**: This indicates PAL's applicability across both STEM and language learning.
3. **Grade-wise analyses (K–10) showed consistently positive gains, with slightly higher effects in secondary grades:** Starting with secondary grades could be more effective — where greater learner autonomy and wider variation in starting proficiency levels probably allow PAL systems to demonstrate stronger impact.
4. **Learning gains were comparable with or without teacher scaffolding, though absence of facilitation showed higher negative variability in impact:** This underscores the need for structured facilitation, by a teacher or basic instructor, to maintain consistency and effectiveness in implementation.
5. **Higher learning gains were observed with increased PAL use up to 90–120 minutes per subject per week, after which the effects plateaued or declined:** This suggests that moderate, regular engagement with ~2 classroom sessions per week per subject can maximise learning outcomes.
6. **The majority of studies (66 of 73) were conducted in school-based settings:** This demonstrates that PAL has proven to be most effective when integrated with in-school settings, as better facilitation and consistent adoption can be ensured.

The Challenge of Heterogeneity in Indian Classrooms

India has demonstrated a robust commitment to educational reform through the National Education Policy (NEP) 2020 and increased budgetary allocations for Samagra Shiksha. Yet, a critical challenge remains while access to schooling has universalized, learning outcomes continue to lag.

Data from the National Achievement Survey (NAS) 2021 and ASER 2023 highlight a persistent crisis - majority of the students are below grade level competency, with deficits widening in higher grades. As per the ASER 2023 report, 43% of Class VI students face challenges in comprehending the main ideas within texts, highlighting significant gaps in reading comprehension skills, and in mathematics, over 50% of learners aged 14–18 struggle with arithmetic tasks expected by Grade 5 students.

A student one grade level behind in Grade 3 may fall an average of four grade levels behind by Grade 8 if left unaddressed. Traditional "one-size-fits-all" instruction is ill-equipped to address this heterogeneity, and smaller group tutoring —while effective—is resource-intensive and unfeasible at scale. Consequently, as indicated by NAS 2021, proficiency gaps widen significantly as students progress: from 39% of students lacking mathematics proficiency in Class 3 to nearly 68% by Class 10.

Personalised Adaptive Learning (PAL): Potential Solution

Personalized and Adaptive Learning (PAL) has emerged as a scalable, technology-led solution to bridge these gaps and tailor education to each learner’s unique needs in real time. However, to understand its efficacy, we must move beyond generic definitions and understand the underlying architecture that powers adaptive interventions.

A robust PAL solution is not merely "digital content"; it is an interconnected ecosystem comprising **three technical-pedagogical models** that create the dynamic intelligence necessary for continuous personalization.

1.1 The Learner Model (The Diagnostic Engine)

The learner model serves as the foundation for adaptation. It goes beyond binary "right/ wrong" scoring to infer the learner's full profile—including prior knowledge, misconceptions, and learning behaviors. It continuously collects signals such as:

- **Response Accuracy:** To identify mastery and persistence.
- **Error Types:** To distinguish between deep conceptual misunderstandings and casual slip-ups.
- **Time on Task:** To gauge engagement levels and confidence.
- **Interaction Logs:** To detect metacognitive behaviors, such as hint usage or self-correction.

1.2 The Domain Model (The Content Map)

This model structures the specific subject content into a **Logically Hierarchical Concept Map**. It links topics across grades (vertical progression) and difficulty levels (horizontal links).

- *Example:* Fractions -> Ratios -> Percentages -> Data Handling.
- By mapping these prerequisite relationships, the Domain Model ensures that the system's adaptivity is guided by curricular logic. It allows the system to recommend targeted remedial loops (e.g., revisiting fractions before attempting ratios) or accelerated pathways for advanced learners.

1.3 The Adaptation Model (The Recommendation System)

This is the "brain" of the operation. The Adaptation Model acts as a Recommendation System (RS) that links the *Learner Model* to the *Domain Model*.

- Using diagnostic inputs (from the Learner Model) and the concept map (from the Domain Model), the RS decides the optimal next step: an easier task, an advanced concept, a visual explanation, or a hint.
- **Operationalizing Adaptivity:** This dynamic adjustment ensures that the content difficulty constantly evolves with the learner, optimizing engagement and preventing frustration.

In the Indian context, this architecture serves as an automated mechanism for **"Teaching at the Right Level" (TaRL)**. It enables the system to identify a Grade 7 student's specific gap in Grade 5 arithmetic and instantly construct a personalized remedial pathway—delivering instructional consistency that is difficult to sustain manually in large classrooms.

Institutionalizing Quality: The Role of EdTech Tulna

EdTech Tulna has played a pivotal role in bringing clarity, rigor, and structure to the PAL ecosystem in India. By partnering with governments, educational authorities, and solution providers, Tulna has ensured that decisions around PAL solutions are evidence-based, scalable, and learner-focused. Central to this work is Tulna's rigorous, transparent, and context-sensitive evaluation rubric, which has empowered stakeholders to assess both the quality and adaptivity of PAL solutions with confidence.

To achieve this, Tulna's focused on:

- **Identifying truly adaptive solutions:** Using 4 criteria and 11 indicators that exclusively focus on assessing the quality of adaptivity, combined with Tulna's proprietary evaluation and sampling approach, ensures that adaptivity is measured accurately and reliably.
- **Ensuring overall quality of PAL:** Adaptivity is just one aspect of a strong PAL product. PALs must also be contextually relevant, comprehensible, aligned to curriculum, and other essential standards outlined in Tulna's framework. By combining these criteria with a review of unique product features, Tulna provides a comprehensive, publicly available view of each product's overall quality.
- **Supporting governments:** **In Haryana**, this framework was adapted as '**Haryana Tulna**' to evaluate PAL solutions aimed at supporting over one million senior school students. Expert committees were trained to apply the evaluation rubric, enabling nuanced, evidence-driven assessments of product adaptivity during procurement. **In Uttar Pradesh**, Tulna's evaluations were embedded into the state's procurement process by NITI Aayog as both a pre-qualification filter and a technical scoring criterion. This approach facilitated the rollout of PAL solutions to **50,400 students across 280 schools**.
- **Beyond government engagement:** Tulna has worked closely with leading PAL providers, helping them enhance their products using research-backed insights. Collectively, these solutions now serve over 5 lakh learners across core subjects from grades K–10, ensuring more adaptive and personalized learning experiences.

Through these efforts, Tulna has not only introduced standardized, evidence-based evaluation tools but has also built capacity among governments to make informed, data-driven decisions and product developers to strengthen their offerings and raise standards. By connecting research, policy, and practice, Tulna has helped transform the PAL landscape in India, ensuring that high-quality adaptive learning solutions reach learners at scale, while continuously improving in effectiveness.

The Need to Go a Step Further : From “Quality” to “Efficacy”

While frameworks like Tulna have successfully defined *what constitutes a high-quality product*, a critical gap remains for decision-makers: understanding *under what conditions* these products deliver the highest return on learning.

Decision-makers today face a practical query: **Not just whether PAL is innovative, but for which grade levels, subjects, and student profiles is it most effective?**

To answer this, EdTech Tulna conducted a **Rapid Evidence Review (RER)** of over **73 PAL implementations worldwide**. This paper moves beyond product features to analyze empirical impact, synthesizing global evidence to provide actionable insights for the Indian education ecosystem.

- Understanding what works, for whom, and under what implementation settings;
- Informing policy and Edtech investment decisions, based on where PAL has delivered significant learning gains, which grade ranges (early grades or secondary levels), and in which subjects (e.g., mathematics, science, language);
- Guiding implementation choices, by helping summarise the impact of key programs in other similar programs (like duration of usage)

This deeper, evidence-informed approach complements the ‘design-quality assessment’ by ensuring that states not only procure the best solutions but also make the right program design choices, strategically aligned with their learner needs, infrastructure, and educational priorities.

Methodology

To enable robust interpretation and cross-study comparison, Tulna developed a consolidated dataset comprising **101 studies** drawn from 20 meta-analyses.

Selection Criteria: Strict filtering was applied to ensure rigor. We selected only studies that:

- Utilized RCT, experimental, or quasi-experimental designs.
- Reported core psychometric information.
- Had an intervention period of four weeks or more.

Exclusions: Outliers (>3 SD), non-school delivery settings, and overlapping grades were excluded to preserve the granularity required.

Final Sample: This yielded a final analytical sample of **73 studies**.

The dataset captured key variables including subject, grade, country, dosage (minutes per subject per week), delivery setting (in-school, after-school, at home), effect sizes (Cohen’s d), sample size, and study duration.

Normalization

Empirical research is often skewed by variables such as alignment of outcome measures with the intervention, the timing of post-intervention assessments, measurement reliability, sample composition, the strength of treatment–control contrasts, and the broader implementation context, including scalability and policy feasibility. While all of these factors are relevant for interpretation, many are insufficiently documented to allow consistent quantitative adjustments.

Two variables, however, emerged as both empirically influential and consistently measurable: sample size and study duration. To correct for this, we applied a **double-normalization** process:

1. **Sample size** affects the stability and replicability of effect estimates. Smaller studies are more prone to sampling volatility, often inflating reported effects, whereas larger, well-powered studies provide more precise and reliable estimates. A scaling factor was applied that increases the weight of studies as sample size grows, but levels off once additional participants contribute little to reliability. This approach reduces the influence of small, volatile studies without giving excessive weight to very large ones.
2. **Study duration** influences the generalizability of results. Short-term studies are more likely to capture "novelty" spikes that may not be sustained, while longer studies provide a clearer indication of durable impact. An exponential decay adjustment was applied to progressively reduce the weight of shorter studies, reflecting the expected attenuation of early gains, with longer-duration studies approaching full weight to signify confidence in lasting effects.

This approach ensures that the findings presented below represent a reliable, bias-adjusted view of PAL's true impact.

Two variables, however, emerged as both empirically influential and consistently measurable: sample size and study duration.

Both adjustments were applied sequentially, first for sample size, then for study duration, to produce a doubly-normalized effect size. This approach enables fair and consistent comparisons across studies of varying designs, contexts, and intervention types. By directly addressing two common sources of bias, the normalization enhances the reliability and interpretability of findings, giving decision-makers a strong evidence-backed foundation for assessing the true impact of PAL interventions.

Key Findings

1. PAL delivered comparable gains across developing and developed world context

To examine PAL's impact across income levels, studies were grouped using the World Bank's 2024 income classification (aligned with WHO conventions). The analysis yielded no statistically significant variance between low-middle income contexts such as India, Pakistan, and Nigeria, upper-middle income countries including China, Turkey, and Russia, and high-income systems like the USA, UK, and Germany. They all reported comparable effect sizes.

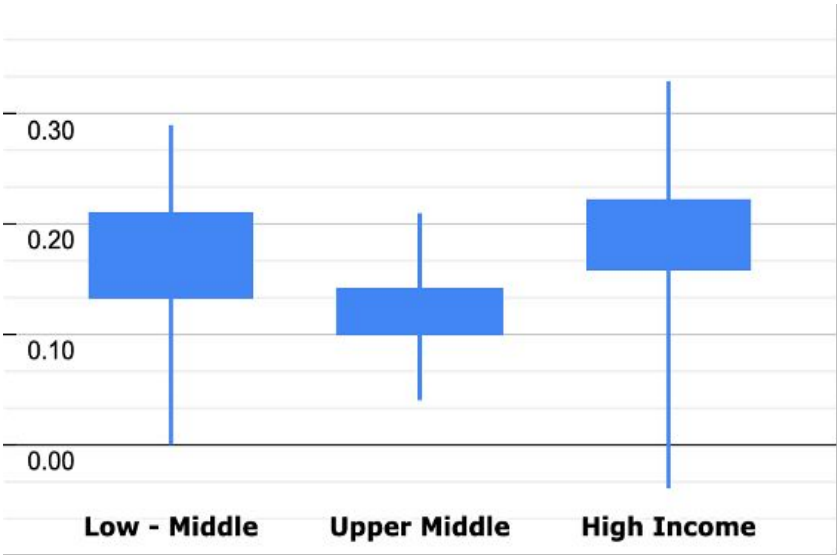

Figure 1: Spread of PAL impact Low-Middle, Upper Middle and High Income countries

This challenges common assumptions that high-income countries achieve substantially higher impact in technological interventions due to stronger infrastructure, better-trained teachers, and better resourced implementation.

Demonstrating that PAL's effectiveness depends primarily on context-specific adaptation and implementation quality, not national income. Well-designed and carefully implemented PAL interventions can therefore deliver consistent learning gains across all economic settings, reinforcing its scalability and universal applicability.

2. Across Maths-English-Science, positive and significant learning gains observed

PAL interventions show positive impact across mathematics, English, and science. While mathematics and English are often highlighted as areas of strength, evidence showed that learners achieve comparable gains in science, demonstrating the approach's versatility across disciplines. Consistent results across subjects reinforce that PAL is not subject-specific but a flexible, scalable tool capable of supporting meaningful learning outcomes in multiple domains.

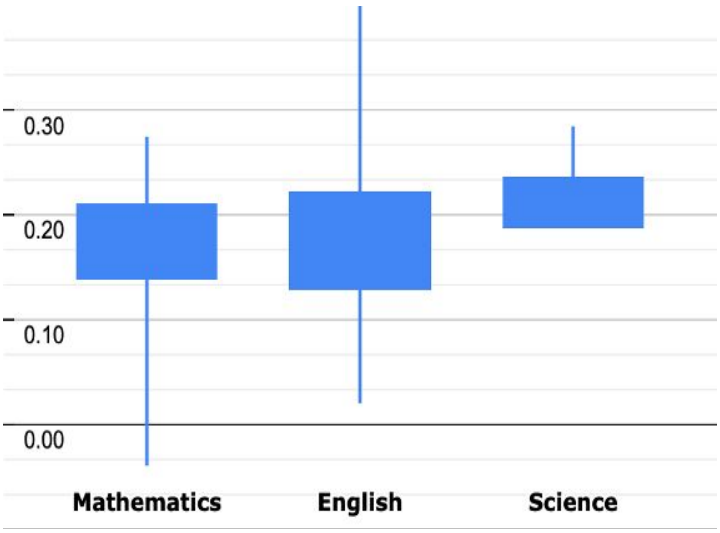

Figure 2: Impact spread across subjects

Note: *There were 25+ studies in Mathematics and English whereas Science had only 4 impact studies but showed most consistent gains with least variability in effect size. This could also be a function of difficult to study, administer studies - so needs to be evaluated further.*

3. Across K-10, PAL showed positive and significant learning gains, stronger in secondary grade

Learning gains are positive across all stages, median effect size of 0.22 in early grades (K–2), 0.21 in grades 3–5, and 0.19 in grades 6–8, the strongest gains are seen in grades 9–10, where the median effect size rises to 0.27. Greater heterogeneity in learning levels in higher grades may have driven higher gains from personalized-adaptive learning. This pattern suggests that PAL can be particularly effective in classrooms with wide learning disparities.

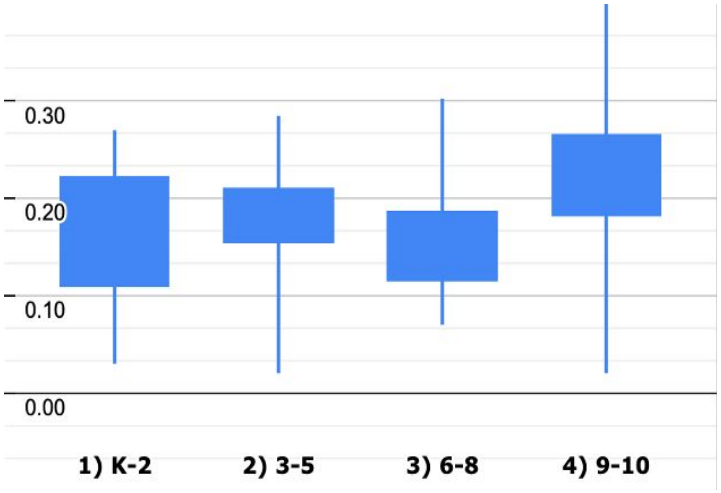

Figure 3: Impact spread across grades

4. Learnings gains comparable with or without Teacher scaffolding, but 'no facilitation' showed higher risk of variable/negative impact

Effect sizes for PAL interventions show broadly comparable results across three levels of facilitation: teacher scaffolding, facilitation, and no facilitation. Each, however, varies in nature. Teacher scaffolding is associated with greater variability in outcomes, with effect sizes ranging from 0.00 to 1.05 and a median of 0.22, pointing to context-dependent impact. Facilitation produces more consistent results, reflected in a narrower distribution and a median effect size of 0.20. By contrast, interventions with no facilitation display higher risk: while the median effect size is 0.21, some fall below zero (minimum -0.04), suggesting that absence of support can at times undermine effectiveness. Together, these findings indicate that learning gains are achievable with or without scaffolding, but lack of facilitation raises the likelihood of variable or even negative outcomes.

Teacher Scaffolding: The role of the teacher is active like choosing the content or activities from possible options provided by the software to meet the learning goals, and/or by providing academic support and feedback.

Facilitation: The role of the teacher or supervisor was limited to providing technical support when supervising the implementation of a programme.

No Facilitation: There is no role of a teacher during implementation.

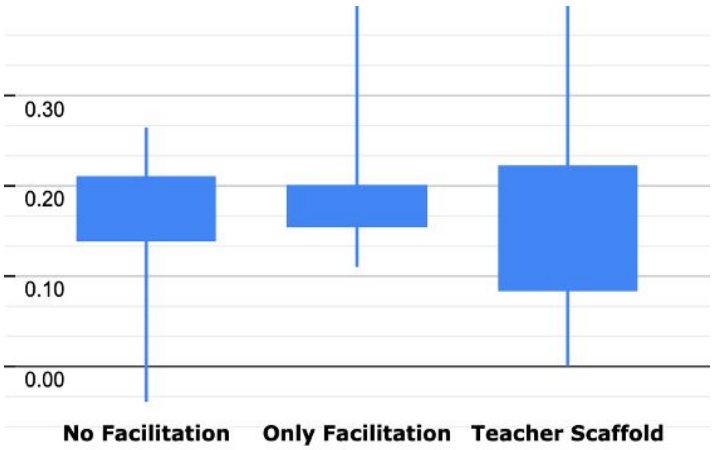

Figure 4: Impact of teacher support across subjects

5. Highest and consistent learning gains observed with 90–150 mins (~ 2-3 periods) of usage per subject per week

The evidence shows strongest and most consistent gains are observed with 90–150 minutes per week, where the distribution of effect sizes is tightly clustered, indicating reliability across contexts/studies. For this band, the median effect size is 0.22. Lower dosages (30–90 minutes) also generate positive results, but with much wider variability: while the median is 0.16, the spread extends from 0.00 to 1.05, suggesting that a few outliers are responsible for inflating the maximum. Higher dosage, 150–240 minutes, shows dispersion and diminishing returns. The median effect size falls to 0.19 and the interquartile range is narrower but shifted lower (0.19–0.29), indicating that the additional time does not substantially improve gains and instead introduces unpredictability.

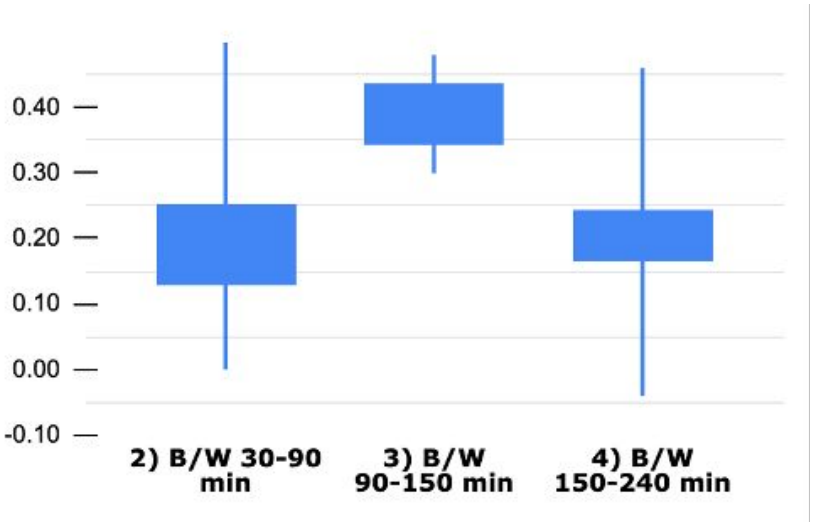

Figure 5: Impact as per dosage

Figure 5 illustrates dosage patterns across all subjects, and a similar trend is observed when analyzed for Maths and English individually (not enough studies for Science). Notably, a larger share of studies per subject cluster within the 90–150 minutes band, emphasizing its consistency and reliability. These findings show that the 90–150 minute range offers the optimal balance, making it the most effective and practical target for weekly engagement per subject.

Together, these findings suggest 90–150 minutes, equivalent to 2–3 periods per subject per week in Indian classrooms, maximizes impact. Too little exposure introduces volatility due to outliers, while too much diminishes average gains.

6. 66 of the 73 studies on PAL interventions were conducted in school setting - showing it can be blended effectively with in-class learning

Most of the studies on PAL interventions were conducted in school, with 66 studies showing a solid average effect (mean 0.25, median 0.22), indicating consistent positive results in this setting. In contrast, there are very few studies on after-school programs (only 2) or combined in-school and after-school programs (5 studies). Because of this small sample size, any insights about these settings are less reliable and harder to draw firm conclusions from. Therefore, the current evidence mainly supports the effectiveness of PAL when implemented during regular school hours.

Recommendations

Based on the synthesis of global evidence, we propose the following guidelines for policymakers and program designers in India and other developing countries.

Optimal Program Design

1. **Dosage:** Ensure **90–120 minutes usage per subject per week**. This translates roughly to at least **2 classroom periods per week**. Even 1 period (30-90 mins) shows significant gains of 0.16 SDs but they increase to 0.22 SDs with 90-150 mins exposure. Consistent and moderate usage of PAL tools appears to be most effective.
2. **Setting:** Prioritize **in-school integration**. Schools provide the structured environment necessary to ensure consistent adoption and mitigate the high dropout rates often seen in purely at-home models. While the potential of at-home or hybrid models is acknowledged, further research is needed to understand how to optimize these approaches, particularly in contexts adult guidance.
3. **Role of Facilitator:** Programs should define a clear "Facilitator" role—whether a subject teacher or a lab instructor—to provide technical oversight and behavioral nudges. This "Human-in-the-Loop" model is essential to prevent the negative variability seen in unsupervised deployments.

Strategic Focus of Programs

1. **Grade Level:** While PAL is effective K-10, states with limited budgets may find the highest immediate ROI in **middle-secondary Grades (6-10)**, where learning gaps are widest and student autonomy is higher.
2. **Low resource contexts:** The evidence confirms that with the right software and implementation design, PAL is as effective in developing contexts as it is in high-income nations.

Conclusion

Global evidence confirms Personalized Adaptive Learning (PAL) is a powerful intervention for addressing learning heterogeneity. However, success relies on two non-negotiable pillars: **product quality** and **program design**. First, not all "adaptive" systems are equal; rigorous standards like **EdTech Tulna** are essential to filter for true pedagogical rigor. Second, implementation must be evidence-led—prioritizing moderate dosage (90–120 minutes/week), strategic targeting (Maths, English, Science, especially in Grades 6–10), and structured facilitation support.

By combining verified, high-quality solutions with these evidence-backed design choices, policymakers can ensure PAL delivers consistent, scalable impact for Indian learners.

EdTech Tulna

Establishing quality standards in EdTech

Developed by

IIT Delhi

IIT Bombay

EdTech Tulna is a public good and is available at no cost

edtech-tulna

contact@edtechtulna.org

www.edtechtulna.org