CHAT2LEARN SUITE

Proof of Concept White Paper

Evidence, Classroom Pilot Results, and Technical Architecture

---

Prepared by: Ariel Kalil, Ph.D. & Michelle Michelini, M.P.P.

Behavioral Insights and Parenting Lab, Harris School of Public Policy, University of Chicago

September 2026

Executive Summary

The Chat2Learn Suite is an AI-powered messaging platform designed to close the early-childhood conversation gap by delivering illustrated, research-informed conversation prompts to teachers and parents. Developed at the University of Chicago's Behavioral Insights and Parenting Lab (BIP Lab), the Suite addresses a well-documented problem: children from low-income families enter school with significantly weaker language skills than their higher-income peers, and the quality of adult–child conversation is the single most powerful lever for closing this gap.

This white paper presents proof-of-concept evidence across three domains: (1) a rigorous classroom needs assessment and a completed three-pilot classroom test demonstrating both the opportunity for a light-touch conversational tool and teachers' ability to use one; (2) published experimental evidence showing that Chat2Learn causally improves the quantity and quality of parent-child language interaction; and (3) the technical architecture powering scalable, real-time prompt delivery across SMS, WhatsApp, and a classroom mobile application.

Key results include a 0.24 to 0.37 SD improvement in children's vocabulary from a six-month randomized controlled trial with 594 low-income families; a lab experiment in which parents using Chat2Learn produced 13% more words, 27% longer sentences, and 92% more open-ended questions; and classroom observations confirming that in 81% of 72 observations, teachers asked closed-ended questions more often than open-ended ones, with 56% of observations revealing missed opportunities for conversation. In classroom usage pilots across 24 pre-K and kindergarten classrooms, teachers used the app on 53.8% of eligible days, delivered 96% of their responses by voice, and roughly doubled their rate of open-ended questions, from 0.23 to 0.49 per minute, while closed-ended questions fell from 0.49 to 0.37 per minute. A randomized comparison within the pilots showed that enhanced implementation support raised usage from 45% to 67% of eligible days. The Chat2Learn Suite is uniquely positioned as the first platform to bridge language-building conversations at school and at home through coordinated, illustrated conversation prompts.

1. The Problem: The Conversation Gap in Early Childhood

1.1 The Science of Early Language Development

Research in developmental science has established that the quality of language input children receive—particularly through interactive, back-and-forth conversation—is the primary engine of early language development (Anderson et al., 2021; Rowe & Snow, 2020). Based on sociocultural theory (Vygotsky, 1978), children learn language through interactions with more knowledgeable others, using the frequency and variety of input to develop representations of words, grammar, and meaning. Since children spend most of their early years at home and in classrooms, the language environments provided by parents and teachers are the dominant influences on language growth.

Language skills at kindergarten entry are among the strongest predictors of later academic achievement and life outcomes (Dickinson & Tabors, 2001; Duncan et al., 1998; Snow et al., 1998). Rowe and Snow (2020) defined three key dimensions of language input that matter most during the preschool years: *interactive features* (open-ended questioning, joint attention), *linguistic complexity* (diverse vocabulary, longer utterances), and *conceptual richness* (decontextualized talk about hypothetical scenarios, explanations, and narratives). Decontextualized language—talk that goes beyond the "here and now" to reference imaginary, past, or future events—is particularly powerful because it requires children to draw on their own knowledge, reason abstractly, and construct meaning without perceptual support (Rowe, 2013; Curenton, 2010).

1.2 Disparities in the Conversation Environment

Yet substantial disparities persist in both the quantity and quality of language input along socioeconomic lines. Children from lower-income families are exposed to fewer conversational turns, less complex vocabulary, and fewer open-ended questions—both at home and at school (Dailey & Bergelson, 2022; Golinkoff et al., 2019; Rowe, 2008; Hoff, 2003). These differences are evident as early as 18 months of age (Fernald et al., 2013) and widen over time, undermining academic achievement and social mobility. Research on school-family communication compounds the problem: a study of a leading school messaging platform found that only 8% of messages between schools and parents were about children's learning—the vast majority focused on logistics such as schedules and reminders.

1.3 Limitations of Existing Approaches

The standard approach to addressing the conversation gap—teacher professional development and parent education programs—suffers from two critical limitations. First, these programs **add to the cognitive and time burden** of the very adults they aim to support without evidence of behavior change that results in better student outcomes. Teachers consistently cite rigid time pressure, limited planning time, and the cognitive demands of classroom management as barriers to richer language interaction. Parents—particularly those in lower-income households managing financial stress, irregular work schedules, and caregiving demands—face similar constraints (Ellwood-Lowe et al., 2022). While teachers and parent alike indicate they would like to improve practices of engaging with children, existing programs are not helping them to close their intention-action gap.

Second, traditional interventions rely on an *explicit* theory of change that explains to adults *why* and *how* to talk with children, assuming that knowledge translates into behavior change. Over the past several decades, numerous interventions have used this approach—from shared book-reading programs to home-visiting interventions to text-based tip-delivery systems like Ready4K and Vroom. While some have shown modest effects, they require sustained engagement with instructional content and depend on parents and teachers independently translating general knowledge into specific conversational behaviors in real time. Research consistently shows that this is extremely difficult for adults challenged by distractions and inattention.

1.4 The Chat2Learn Theory of Change

The Chat2Learn Suite takes a fundamentally different approach. Grounded in cognitive load theory (Sweller, 2011; 2020), it operates through an *implicit* theory of change: rather than teaching adults what to say or why to say it, the platform directly supplies the "what to say." By providing ready-to-use, open-ended conversation prompts paired with engaging illustrations, Chat2Learn reduces the cognitive burden of initiating high-quality conversations and functions as just-in-time behavioral support that fits seamlessly into existing routines.

The Suite is designed around three core principles. First, it **targets adult behavior change**, reflecting the developmental science that adult–child interaction is the primary engine of early learning. Second, it **reduces cognitive load** for both teachers and parents by supplying the conversational content itself, not just advice about what types of conversations to have. Third, it **bridges school and home** by coordinating the delivery of synergistic, illustrated prompts to teachers and parents—a design feature focused on extending and deepening classroom conversations at home – which no existing product offers.

2. Classroom Proof of Concept: Needs Assessment and Pilot Results

With support from a 2024 Institute of Education Sciences (IES) Seedlings to Scale grant, the Chat2Learn team conducted a comprehensive needs assessment in PreK and Kindergarten classrooms across low-income communities in Chicago. The assessment was designed to answer three questions: (1) Is there a demonstrable need for a tool like Chat2Learn in classrooms? (2) Do teachers perceive this need? (3) What are teachers' preferences for program use?

2.1 Study Design

The research team completed **72 structured observations** (40 whole-group, 12 small-group, and 20 one-on-one interactions) across **17 classrooms** (10 PreK, 7 Kindergarten). Observers used a standardized rubric to code the frequency and type of teacher-initiated questions, whether teachers took advantage of opportunities to engage students in conversation, and classroom contexts where a digital conversation aid might be useful. The team also conducted **21 semi-structured interviews** with teachers (13 PreK, 8 Kindergarten) to understand their beliefs, practices, and perceived barriers.

2.2 Key Findings

Finding 1: Teachers predominantly ask closed-ended questions

Across all 72 observations, teachers asked an average of 1.2 questions per minute. However, closed-ended questions dominated: they were the most common question type in **81% of observations** (58 of 72). Teachers averaged 0.89 closed-ended questions per minute compared to only 0.31 open-ended questions per minute. In many observations, teachers asked **no open-ended questions at all**.

Finding 2: Teachers frequently miss opportunities for meaningful conversation

Observers assessed whether teachers took advantage of natural opportunities to engage students in conversation (e.g., responding to a child in a way that invites continued engagement, asking a question rather than making a comment, or drawing a disengaged student into dialogue). In **56% of observations, teachers left opportunities on the table.** Only 19% of observations were coded as "Always" taking advantage of opportunities. PreK teachers were particularly likely to miss opportunities: only 11% of PreK observations were rated "Always" compared to 27% in Kindergarten.

Table 1. Observer Ratings: "Did the teacher take advantage of opportunities to engage students in conversation?"

Rating	All Classrooms	PreK Only	K Only
Always	19%	11%	27%
Often	25%	29%	22%
Sometimes	22%	26%	19%
Rarely	25%	29%	22%
Not at all	8%	6%	11%

Source: IES Seedlings to Scale Needs Assessment, n = 72 observations across 17 classrooms (2025–2026).

Finding 3: Teachers see the need and want a tool like Chat2Learn

All 21 interviewed teachers affirmed that engaging children in conversation is important for building language skills. Almost half reported they were **not spending as much time in conversation with students as they would like.** Notably, the three teachers who performed best on observer ratings of engaging students in conversation all self-reported that they were not spending enough time in conversation—suggesting a pervasive intention–action gap even among the most skilled educators.

Teachers identified several barriers to effective conversation:

- **Time constraints:** Rigid curricula and prescribed schedules leave little room for organic conversation. As one teacher reported, "Time is obviously the biggest challenge. When would I find the time when I'm already feeling so rushed?"
- **Diverse language abilities:** Managing classrooms with children at widely varying language levels makes it difficult to calibrate questions appropriately in real time.
- **Behavior management demands:** Teachers' attention is frequently pulled toward managing disruptive behavior, leaving less cognitive bandwidth for initiating thoughtful dialogue.
- **Finding engaging topics:** Several teachers noted difficulty generating novel, age-appropriate topics for on-the-spot conversation, especially during unstructured activities such as transitions and free play.

Critically, when shown a demonstration of the Chat2Learn classroom app, **100% of interviewed teachers expressed interest** in using it; no teachers expressed disinterest. Teachers identified multiple use cases: whole-group instruction, individual interactions, centers and free play, structured circle time, and transitions. Several teachers specifically noted the value of using Chat2Learn prompts during times they found most challenging, such as transition periods and behavior management moments. One teacher

explained, "I'd throw out a question during transition time, when we are moving from a small group activity to recess or when a child is waiting in line."

Teachers expressed particular enthusiasm for two features: the app's **accompanying illustrations** (which they viewed as engaging visual anchors for conversation) and the **caregiver engagement component** (the ability for parents to receive related prompts at home). Multiple teachers independently identified parent engagement as a high-value use case, with one noting, "I am always trying to get my parents to understand the importance of conversations, but I know that they aren't talking with their children to the extent we would want them to; this would be a good resource."

Finding 4: Observers identified specific classroom contexts where Chat2Learn would be most valuable

When asked to identify specific moments during their observations where a digital conversation aid would have been useful, observers documented opportunities across a range of classroom contexts:

- **Circle time and morning meetings:** Prompts could introduce topics or serve as attention-grabbing conversation starters connected to calendar themes (seasons, holidays, community events).
- **Curriculum additions:** Prompts could supplement reading and language lessons with open-ended follow-up questions or serve as turn-and-talk prompts for small-group discussions.
- **Behavior management and transitions:** Prompts could productively occupy disengaged or disruptive students and support smoother transitions between activities.
- **Centers and open play:** Chat2Learn could function as a dedicated classroom center activity or pair with existing activities (e.g., asking questions relevant to a drawing or building activity).
- **Snack and mealtimes:** Informal settings where conversation happens naturally but often defaults to simple directives.

For 17 of approximately 30 teachers observed, observers clearly indicated that there were opportunities where a digital conversation aid would have supported richer teacher–student dialogue. Some observers also noted potential barriers in specific scenarios (e.g., preserving solo play time, or classrooms where the teacher already excelled at organic back-and-forth), reinforcing that the tool should be designed as a flexible complement to existing practice rather than a prescribed protocol.

2.3 Classroom Pilot Results

Building on the needs assessment, the team conducted three sequential usage pilot tests of the Chat2Learn classroom app, each lasting three weeks, across 24 pre-K and kindergarten classrooms in Chicago partner networks. Each participating teacher received an iPad loaded only with Chat2Learn and locked to prevent other downloads, isolating use of the tool from other digital activity. The team collected backend usage data, pre- and post-pilot structured classroom observations, and teacher and parent surveys.

Teachers used the app on 53.8% of eligible days, opening an average of 22.5 prompts and responding to 61.9% of them. Notably, 96% of responses were dictated by voice rather than typed, a strong signal that

dictation is how teachers will use a tool of this kind during instruction. Pre- and post-pilot structured observations showed open-ended questions rising from 0.23 to 0.49 per minute while closed-ended questions fell from 0.49 to 0.37 per minute. Because pilot access was not randomized and the observations were not fully paired, these movements are descriptive and consistent with the theory of change rather than causal. Teachers integrated the tool into routines that already existed, including morning meetings, transitions, free play, small groups, English-language-learner support, and one-on-one work, which matters as much as the usage rate itself.

One of the three pilots included a randomized test of implementation support. All teachers received a live 30-minute onboarding demonstration; enhanced-support teachers additionally received a hands-on demonstration and reminders and offers of help three times weekly, against once weekly for the low-support group. Enhanced-support teachers used the app on 66.8% of eligible days compared with 44.6% under low support, and responded to 72.7% of prompts compared with 54.2%. This is the one causal finding from the usage pilots, and it establishes that implementation support is not optional.

At endline, 14 of 23 teachers agreed that Chat2Learn enhanced classroom conversations and 10 of 23 that it enhanced student learning. Among parents, 26 of 39 agreed that the prompts helped them have better conversations with their children. Teachers consistently described strong child enthusiasm for the prompts.

Table 2. Usage pilot implementation metrics and randomized support comparison

Measure	Overall (N=24)	Enhanced (N=10)	Low (N=14)
Eligible testing days	14.0 (1.1)	14.0 (0.7)	14.1 (1.4)
Days app used	8.1 (2.9)	9.7 (2.3)	7.0 (2.8)
% eligible days app used	53.8 (22.4)	66.8 (12.5)	44.6 (23.6)
Prompts opened	22.5 (15.2)	22.9 (19.0)	22.1 (12.6)
% prompts responded to	61.9 (22.0)	72.7 (16.8)	54.2 (22.5)
% responses by voice	96.2 (11.4)	99.5 (1.5)	93.6 (14.8)

Note. Cells are teacher-level means (standard deviations) within each teacher's eligible testing window. Teachers were randomized within each sequential pilot. Observed group means are reported; inferential analysis is ongoing.

2.4 What Usage Piloting Showed Is Not Yet Ready

Piloting was equally informative about the tool's limits. Intent to continue was mixed, and 21 of 23 teachers reported at least one challenge in using the tool. Technical reliability was the weakest dimension, with teachers encountering login failures, slow loading, audio recording problems, language-switching errors, and prompts that did not refresh. AI-generated follow-ups were sometimes mismatched to a child's age, home language, or expressed topic of interest. School and district leaders added two procurement conditions: content aligned to the curriculum a school already uses, and clear, usable metrics on program use and impact. Cost-effective professional development was also named a priority.

These findings define five areas of development required before a rigorous impact evaluation: (1) technical reliability; (2) alignment of prompt content with grade-level curricula; (3) scalable teacher onboarding and professional development; (4) an educator-facing dashboard reporting metrics educators can act on; and (5) direct delivery of prompts from teachers to families with a structured channel for family responses. In these initial usage pilots, the research team delivered family prompts manually, so the school-home link has not yet been tested as a product feature.

3. Published Evidence: Experimental Results on Parent–Child Interaction

While the classroom component is in active development, the parent-facing side of Chat2Learn has been rigorously evaluated through multiple experimental studies. This section summarizes evidence from a lab-based proof-of-concept experiment (Lu & Kalil, 2025) and a large-scale randomized controlled trial (Rury, Kalil, Mayer, & Bresciani Andaluz, 2025).

3.1 Lab Experiment: Causal Impact on Parent Language Behavior (Lu & Kalil, 2025)

Lu and Kalil (2025) conducted a controlled experiment with 63 parent–child dyads recruited from Head Start and state-subsidized preschool centers serving low-income communities. Of the 63 participating parents, 87% were mothers. The children (65% girls) ranged in age from 3 to 6 years. Among families who reported educational attainment, 49% had a high school diploma or less, 34% had some college, and 17% held a bachelor's degree or higher. The racial/ethnic composition was 65% Hispanic/Latino, 19% Black/African American, 5% White, and 5% other; 6% did not report race/ethnicity.

Dyads were randomly assigned to either a control condition (10 minutes of free interaction time in a quiet room) or a Chat2Learn condition (10 minutes of free time with access to an iPad preloaded with eight illustrated, decontextualized conversation prompts). Example prompts included, "What could we do if we had eight arms like an octopus?" and "If you could have one superpower, what would it be?"—each paired with a hand-illustrated image. Critically, parents in the Chat2Learn condition were *not instructed to use the prompts*; they chose to do so voluntarily, and 100% of intervention parents used them. All sessions were videotaped and coded by trained research assistants who were blind to the study hypotheses and conditions.

Coding captured three dimensions of interaction: (1) *moment-to-moment parent engagement*, using continuous second-by-second state coding of joint engagement, non-digital disengagement, and digital disengagement; (2) *quantity and quality of parental language input*, including total words, utterances, unique words, mean length of utterance, and open-ended questions; and (3) *child language output*, using the same linguistic measures. Inter-rater reliability was established through independent double-coding, with intra-class correlations ranging from .82 to .99.

Results: Parent Engagement

Parents in the Chat2Learn condition spent marginally more time in **joint engagement** with their children: 92% of session time (SD = 17%) compared to 78% (SD = 30%) in the control group ($p = .07$).

Non-digital disengagement dropped significantly, from 9% in the control group to just 2% in the experimental group (p = .05). Notably, all parents in the intervention group *chose to use the prompts voluntarily*—they were not instructed to do so.

Results: Parent Language Input

Multivariate analysis controlling for child age, home language, and session time revealed highly significant effects on parental language quality:

Table 3. Treatment Effects on Parent Language Input (Lu & Kalil, 2025)

Measure	Control M (SD)	C2L M (SD)	Significance
Total words	551 (308)	624 (203)	p = .01
Mean length of utterance	3.54 (0.75)	4.50 (0.83)	p < .001, η_p^2 = .28
Open-ended questions	19.52 (15.43)	37.43 (12.37)	p < .001, η_p^2 = .32
Unique words	179 (70)	191 (45)	p = .05
Joint engagement (% time)	78% (30)	92% (17)	p = .07

Source: Lu & Kalil (2025). N = 63 parent–child dyads. All analyses control for child age, home language, and session time.

In practical terms, parents using Chat2Learn produced **13% more total words**, sentences that were **27% longer and more complex**, and **92% more open-ended questions.** The nature of the conversations also shifted dramatically: while control-group interactions focused on the immediate "here and now" (naming objects, commenting on surroundings), Chat2Learn parents engaged in rich *decontextualized talk*—hypothetical scenarios, explanations, and opinion-sharing—the type of language interaction most strongly associated with vocabulary growth and school readiness (Rowe, 2013).

Children's own language output did not differ between conditions (all p's > .10). Lu and Kalil note that decontextualized talk is demanding for preschoolers and suggest adaptive follow-up prompts to help parents sustain these exchanges.

Why Chat2Learn Works: The Cognitive Load Mechanism

A key finding from the study is that even in a controlled setting with no competing demands—no siblings, no household tasks—parents in the control group still showed substantial disengagement (averaging 22% of session time disengaged, including 13% on their phones, with some dyads fully disengaged for the entire session). This pattern highlights the challenge of sustaining parent–child interaction even under ideal conditions.

Ecological models of parenting suggest that parents' behaviors are shaped by culturally rooted belief systems about child development. But beliefs seem unlikely to explain the disengagement seen here: 79% of parents in the study rated talking with their children as extremely important (10 out of 10) for supporting language development. Prior research has documented that parents—especially those experiencing financial stress—report fatigue and difficulty generating new conversation topics (Ellwood-Lowe et al., 2022). Parents in this study reported the same barriers: 39% named fatigue as a barrier to conversation at least one or two days a week, and 39% said they struggled to come up with new topics. By offering concrete, ready-to-use prompts, Chat2Learn may lessen this cognitive load and help parents

stay engaged. The mechanism is not information delivery (teaching parents *why* to talk) but behavioral scaffolding (giving parents *what* to say in the moment).

Post-study surveys provided additional insight: parents in the Chat2Learn condition reported that the prompts gave them **new ideas for conversations** and helped them understand **which questions to ask.** Several parents described unexpected bonding moments. In one family, a child's answer to "If you could fly, where would you go?" revealed a wish to visit cousins in Mexico. The answer surprised the parent and sparked a richer conversation about family. Another parent said, "It gave us something to talk about together." This suggests that even an implicit intervention may shift parental knowledge and self-efficacy over time—a secondary benefit beyond the immediate behavioral effects.

3.2 At-Scale RCT: Impact on Children's Vocabulary (Rury, Kalil, Mayer, & Bresciani Andaluz, 2025)

A separate large-scale randomized controlled trial evaluated a static text-message-based version of the Chat2Learn parent program with **600 low-income families** (Rury, Kalil, Mayer, & Bresciani Andaluz, 2025). Families were randomized to receive one of two treatments that sent parents 3–5 weekly conversation prompts, each paired with an illustrated image, or to a control condition. Children's vocabulary was assessed using standardized measures at baseline and follow-up.

The trial demonstrated a treatment effect of **0.37 standard deviations on children's vocabulary**, equivalent to approximately **eight months of developmental growth.** This effect size is remarkable in the early-childhood intervention literature, where meta-analyses of low-cost parent interventions typically report effects of 0.15–0.25 SD (Roberts et al., 2019). The magnitude of the Chat2Learn effect is comparable to that of far more intensive and expensive interventions, such as home-visit programs and multi-week shared-book-reading curricula—yet Chat2Learn achieves this effect through a fully automated, text-message-based delivery system that requires no in-person contact with families.

The RCT results provide strong evidence that the Chat2Learn mechanism—reducing cognitive load by providing conversational content rather than educating parents about conversational practices—can produce sustained, meaningful gains in children's actual learning outcomes, not just changes in parent behavior.

3.3 Scaling and Engagement Metrics

The Chat2Learn parent program has now been deployed at scale, reaching over **5,000 families** across Illinois, with international pilots in Peru and India via WhatsApp. Engagement metrics are summarized below:

Metric	Chat2Learn Performance
Parent retention rate	80%
Parent reply rate	50%
Industry average reply rate (text programs)	10–15%
Families reached (Illinois)	5,000+

International reach	Peru and India (via WhatsApp)

The 50% parent reply rate—three to five times the industry average for text-based programs—is likely to be an underestimate of usage as parents who never reply to the program may be using and benefitting from the prompts delivered to them 3-5 times weekly with their children.

4. Technical Architecture and Platform Design

The Chat2Learn Suite is built on a modern, cloud-native technology stack designed for scalability, real-time AI-powered interactions, and multi-channel delivery. This section describes the platform architecture, data schema, and deployment infrastructure.

4.1 Architecture Overview

The platform consists of five integrated components:

1. **Quart/Flask Application (Python):** The core backend handles all incoming requests, processes data, and manages interactions with downstream services. Quart's async capabilities ensure high performance under concurrent load.
2. **OpenAI/ChatGPT Integration:** Powers the AI layer for generating real-time, individualized follow-up questions based on each child's responses. The LLM is fine-tuned on a researcher- and artist-produced dataset of conversation prompts and illustrations, keeping conversations within the child's zone of proximal development.
3. **Firebase (Database):** Manages user profiles, conversation logs, child data, daily questions, and engagement streaks. Firestore's real-time syncing enables immediate updates across the teacher app and parent SMS/WhatsApp channels.
4. **Google Cloud Storage:** Hosts the library of hand-illustrated images that accompany each conversation prompt, served via secure URLs.
5. **Twilio (SMS/WhatsApp):** Manages all outbound messaging to parents, supporting both SMS for U.S.-based families and WhatsApp for international deployments. Message delivery records are maintained in Firebase for auditing and analytics.

The entire system is deployed on **Google Cloud Platform (GCP)** using Cloud Run for auto-scaling containerized deployments, with CI/CD pipelines via GitHub Actions for automated testing, linting, and staged rollouts across development, staging, and production environments.

4.2 Data Schema

The Firebase database is organized around five core entities designed to support both the parent and classroom programs:

Entity	Description

Users	Parent/teacher profiles including language preference, phone number, notification settings, engagement streak data, and data-sharing consent.
Children	Child records linked to parent/teacher users, including name, birthday, and profile information for age-appropriate prompt calibration.
Conversations	Complete interaction logs storing each message exchange (content, role, timestamps, transcripts, voice notes, and images), linked to both user and child records.
Daily Questions	The prompt library organized by date, with bilingual support (English/Spanish), keyword tags, and linked illustration URLs.
Images	Hand-illustrated image assets with descriptions and semantic tags, enabling prompt–illustration matching.

The relational structure—Users → Children → Conversations, with Daily Questions and Images as shared resources—supports both the parent SMS/WhatsApp flow and the classroom tablet app through a unified API. This architecture enables the **linked Suite model**: when a teacher uses a prompt in class, an extension prompt can be delivered to parents that evening, creating a reinforcing loop between school and home learning.

4.3 Multi-Channel Delivery

The platform's multi-modal API currently supports three delivery channels:

- **Classroom Tablet App:** A native mobile application (currently in TestFlight) that delivers on-demand prompts with illustrations, enabling teachers to initiate conversation during any classroom activity. The app is designed for minimal interaction—a teacher can pull up a prompt and its accompanying illustration in seconds, then put the device down to focus on the conversation.
- **SMS (via Twilio):** Delivers 3–5 weekly prompts to U.S.-based parents with inline illustrations and real-time, on-demand follow-up questions powered by the LLM. The system tracks message delivery, parent replies, and conversation completion for engagement analytics.
- **WhatsApp (via Twilio):** Identical functionality for international deployments, currently serving families in Peru and India with multi-lingual prompts.

4.4 AI-Powered Personalization

A core differentiator of the Chat2Learn platform is its use of a fine-tuned large language model to generate real-time, individualized follow-up questions. When a parent texts back their child's response to a prompt, the LLM generates a contextually appropriate follow-up that keeps the conversation within the child's *zone of proximal development* (Vygotsky, 1978)—challenging enough to promote growth, but accessible enough to sustain engagement.

The LLM is trained on a proprietary dataset of researcher-produced and artist-illustrated conversation prompts, ensuring that generated content adheres to evidence-based principles of decontextualized talk, open-ended questioning, and age-appropriate language. This approach addresses a limitation

identified in early user testing: without adaptive follow-up, initial prompts alone may not be sufficient to sustain extended dialogue, particularly for younger children or dual-language learners who may need additional scaffolding.

AI translation powers two-way communication in the user's language of choice. The AI agent responds in the language the user inputs. In the teacher mobile app, audio memos recorded by the teacher are auto-transcribed and responded to in real-time by the AI chatbot in the language being spoken. The program LLM recognizes and can respond to users in over 50 widely spoken languages.

4.5 Security and Compliance

Given the sensitivity of data involving children, the platform implements robust security practices. All Twilio and OpenAI credentials are stored in secure environment variables rather than hard-coded. Firebase Security Rules control data access at the document level. Google Cloud IAM manages service-to-service permissions. The system maintains complete audit trails of all message delivery and conversation data, supporting both research integrity requirements (for IRB-approved studies) and future FERPA/COPPA compliance needs as the product moves toward school-district deployment.

5. Recognition, Funding, and Partnerships

The Chat2Learn Suite's development has been supported by competitive grants and awards that validate both the scientific rigor and commercial potential of the platform:

- **Becker Friedman Institute for Economics at the University of Chicago (2025):** Funded the proof-of-concept lab experiments with parent-child dyads.
- **Kiphart Center for Global Health and Social Development at the University of Chicago (2025):** Funded the adaptation and pilot RCT of the parent program in Peru.
- **Institute of Education Sciences (IES) Seedlings to Scale Grant (2024):** Funded the Phase I classroom development, needs assessment, and pilot testing program described in this paper. A Phase II application for full development and a randomized impact evaluation was submitted in September 2026.
- **MIT Solve & Bill & Melinda Gates Foundation AI-Assisted Assessment Challenge Prize (2024):** Recognizes the Chat2Learn Suite as a winning solution for AI-powered educational assessment and intervention.
- **J-PAL at MIT Research Grant (2022, 2024):** Funded the parent-program RCT demonstrating the 0.37 SD vocabulary effect and a current pilot RCT of the AI-integrated parent program.
- **Robert R. McCormick Foundation Implementation Grant (2023):** Supported scaling of the parent program across Illinois.
- **Illinois State Board of Education Grant (2020):** Funded the initial parent program development and pilot deployment.

6. Competitive Positioning

The early-childhood EdTech landscape includes several products that address either parent engagement or classroom communication, but none that bridge both through coordinated conversation prompts. The Chat2Learn Suite's competitive advantages are summarized below:

Dimension	Competitors	Chat2Learn Suite
Theory of change	Information delivery: tell parents/teachers what to do (Ready4K, Vroom, ReadyRosie)	Behavioral nudge: supply the "what to say" to reduce cognitive load and change behavior directly
School–home linkage	None: parent tools (Ready4K, Vroom) operate independently; school tools (ClassDojo, TalkingPoints) enabling messaging focus on logistics	Coordinated: synergistic prompts delivered to teacher and parent, creating a reinforcing loop; sharing conversation snapshots build partnership
AI personalization	Static content libraries; no real-time adaptation to child responses; use of AI to provide/summarize information	LLM-powered follow-up questions adapt in real-time to each child's responses and developmental level; use of AI to augment connection
Evidence base	Correlational, quasi-experimental evidence (if any); one RCT (Ready4K) shows 0.11 SD effect on a constrained literacy skill	RCT-validated: 0.24–0.37 SD vocabulary effect with 594 families; lab-validated parent behavior change; classroom pilots in 24 pre-K and kindergarten classrooms showing open-ended questioning roughly doubling, plus a randomized test showing implementation support raises usage from 45% to 67% of days

7. Conclusion and Path Forward

The evidence presented in this white paper demonstrates that the Chat2Learn Suite addresses a well-documented, high-impact problem with a solution that is scientifically grounded, technically scalable, and now tested in real classrooms. The needs assessment confirms demand for a light-touch conversational tool; the completed classroom pilots show that teachers will use it inside existing routines and that their questioning shifts in the predicted direction; the randomized support test shows what implementation requires; and prior experimental evidence establishes that the Chat2Learn approach causally improves parent language behavior and children's vocabulary. What initial usage piloting could not establish is causal impact on classroom practice or on children, because pilot access was not randomized and no child assessments were collected. That is the work of the next phase.

With additional investment in the US and globally, the Chat2Learn Suite will:

1. Launch and move beyond the MVP of the linked Suite (classroom + parent) by completing the five areas of development usage piloting identified: technical reliability, curriculum-aligned content, scalable onboarding and professional development, an educator-facing dashboard, and direct school-home prompt delivery with structured family responses.

2. Conduct a rigorous impact evaluation of the completed Suite. A two-arm cluster-randomized trial in 60 pre-K and kindergarten classrooms serving approximately 1,200 children will test effects on the quality of teacher-child language interaction and on children's expressive language.

3. Scale through school, NGO, and government networks in the US and globally, leveraging high-impact partnerships and the platform's multilingual, multi-channel delivery capability, with recurring costs projected below $200 per teacher per year.

4. Incorporate as an independent entity with a dedicated founding product team positioned to pursue a B2B revenue model with government and school district contracts while maintaining the mission-driven focus on equity.

The Chat2Learn Suite represents a rare convergence of rigorous developmental science, behavioral economics, and scalable AI technology—all in service of a simple, powerful idea: that giving adults the right words at the right moment can transform the conversations that shape children's futures.

References

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. Child Development, 92(2), 484–501.

Dailey, S., & Bergelson, E. (2022). Language input to infants of different socioeconomic statuses: A quantitative meta-analysis. Developmental Science, 25(3), e13192.

Dickinson, D. K., & Tabors, P. O. (2001). Beginning literacy with language: Young children learning at home and school. Paul H. Brookes Publishing.

Duncan, G. J., Yeung, W. J., Brooks-Gunn, J., & Smith, J. R. (1998). How much does childhood poverty affect the life chances of children? American Sociological Review, 63(3), 406–423.

Ellwood-Lowe, M. E., Foushee, R., & Srinivasan, M. (2022). What causes the word gap? Financial concerns may systematically suppress child-directed speech. Developmental Science, 25(1), e13151.

Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. Child Development, 90(3), 985–992.

Lu, L., & Kalil, A. (2025). Chat2Learn: A proof-of-concept evaluation of a technology-based tool to enhance parent–child language interaction (BFI Working Paper No. 2025-119). University of Chicago.

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. Journal of Child Language, 35(1), 185–205.

Rowe, M. L. (2013). Decontextualized language input and preschoolers' vocabulary development. Seminars in Speech and Language, 34(4), 260–266.

Rowe, M. L., & Snow, C. E. (2020). Analyzing input quality along three dimensions: Interactive, linguistic, and conceptual. Journal of Child Language, 47(1), 5–21.

Rury, D., Kalil, A., Mayer, S., & Bresciani Andaluz, D. (2025). Talking about words boosts preschool-age children's vocabulary: Evidence from a parent intervention (BFI Working Paper No. 2025-24). University of Chicago.

Sweller, J. (2011). Cognitive load theory. In Psychology of Learning and Motivation (Vol. 55, pp. 37–76). Academic Press.

Sweller, J. (2020). Cognitive load theory and educational technology. Educational Technology Research and Development, 68(1), 1–16.