NBER WORKING PAPER SERIES

BELIEFS, ATTENTION, AND INVESTMENTS IN EARLY CHILDHOOD

Flavio Cunha
Snejana Nihtianova
Jessica Rood
Anja-Lize van der Merwe

Working Paper 35150
http://www.nber.org/papers/w35150

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2026

Financial support from the Smith Richardson Foundation, Powell Foundation, Brown Foundation, Samuels Family Foundation, and Episcopal Health Foundation is gratefully acknowledged. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Beliefs, Attention, and Investments in Early Childhood
Flavio Cunha, Snejana Nihtianova, Jessica Rood, and Anja-Lize van der Merwe
NBER Working Paper No. 35150
April 2026
JEL No. C93, D83, D91, I20, J13

ABSTRACT

We develop a model of parental investment in early childhood in which bandwidth-constrained parents hold distorted beliefs about the returns to responsive interaction. When capacity falls short of aspiration, motivated reasoning provides relief: the parent distorts her working belief downward, rationalizing low engagement. Distorted beliefs suppress responsive parenting, which generates no calibration evidence, which, in turn, confirms the distortion. The model identifies four channels through which interventions escape this self-sealing trap. A randomized controlled trial confirms the model's predictions on parental beliefs, measures of responsive parenting, behavioral engagement, and absence of impact on materials or placebo outcomes.

Flavio Cunha
Rice University
Department of Economics
and NBER
flaviocunha@rice.edu

Snejana Nihtianova
Texas Policy Lab, Rice University
snejana.nihtianova@rice.edu

Jessica Rood
Rice University
jr120@rice.edu

Anja-Lize van der Merwe
Rice University
anjalize.vdm@rice.edu

An Online Appendix is available at http://www.nber.org/data-appendix/w35150

1 Introduction

The economics literature on early childhood development models parental investment as a process in which parents lead and children follow. In the technology-of-skill-formation framework, the child is a passive recipient: the parent chooses how much to invest, and the production function maps those investments into developmental outcomes (see, e.g., Cunha and Heckman 2007; Almond and Currie 2011; Cobb-Clark, Salamanca, and Zhu 2019). More recent models endow the child with agency, but preserve the sequential structure: the parent's decisions come first, and the child follows with a response (Akabayashi 2006; Lizzeri and Siniscalchi 2008; Cosconati 2008; Doepke and Zilibotti 2017; Seror 2022; Del Boca, Flinn, Verriest, and Wiswall 2026; Verriest 2024). This paper reverses the timing within the interaction episode. The child signals first, by emitting a vocalization, a gesture, a bid for joint attention, and the parent's investment is a response to that signal. The developmental return to the parent's response is increasing in both the intensity of the child's cue and the degree to which the parent's response is calibrated to that cue, a prediction consistent with a large developmental literature establishing that it is the contingent, cue-calibrated nature of parental responses, not the mere quantity of parental input, that drives early development (Ainsworth, Blehar, Waters, and Wall 1978; Ainsworth 1985; Landry, Smith, and Swank 2006).

Responsive parenting requires the parent to navigate two sequential barriers. The first is **believing**: the parent must hold accurate beliefs about the returns to responsive interaction (Cunha, Elo, and Culhane 2013; Boneva and Rauh 2018; Attanasio, Cunha, and Jervis 2019). The second is **acting**: even a parent who believes must allocate cognitive attention to reading child cues, which is costly when bandwidth is constrained by poverty, stress, and competing demands (Shah, Mullainathan, and Shafir 2012; Mani, Mullainathan, Shafir, and Zhao 2013). Low bandwidth makes both barriers harder simultaneously: it raises the psychological cost of accurate beliefs, because the parent knows she cannot fully deliver on them, and it degrades the quality of cue detection, because attention is expensive. When both barriers are active, they interact to produce a self-sealing trap. The distorted beliefs suppress responsive engagement, uniform responses generate no calibrating evidence, and the distortion is confirmed.

We develop a two-stage model that formalizes both barriers. In Stage 1, the parent chooses a working belief about the returns to responsive interaction, trading off the material value of accurate beliefs against the psychological cost of holding beliefs she cannot act on (Bénabou and Tirole 2002; Bénabou 2015; Hyland and Albarracin 2025). In Stage 2, taking the working belief as given, the parent decides how attentively to monitor child cues, using the rational inattention framework of Sims (2003) with Shannon mutual information costs. The model identifies four channels through which interventions can escape the trap: the *belief channel*, which raises the accuracy cost of downward distortion; the *cue-detection expertise channel*, which lowers the threshold belief required for cue detection to activate; the *bandwidth channel*, which reduces the attention cost and the psychological cost simultaneously; and the *trap compression channel*, which shrinks the set of prior means for which the trap is structurally active.

The economics literature on parenting has extensively formalized the control dimension of parental behavior, meaning the degree to which parents direct, monitor, encourage, incentivize, and even discipline their children (Akabayashi 2006; Cosconati 2008; Doepke and Zilibotti 2017; Lizzeri and Siniscalchi 2008; Del Boca, Flinn, Verriest, and Wiswall 2026; Verriest 2024). This paper formalizes the responsiveness dimension, meaning the degree to which parents read and respond to child-initiated communicative cues, which the developmental psychology literature identifies as a consequential dimension for early development (Ainsworth, Blehar, Waters, and Wall 1978; Ainsworth 1985; Landry, Smith, and Swank 2006). The closest antecedents are Lizzeri and Siniscalchi (2008), Seror (2022), and Cobb-Clark, Salamanca, and Zhu (2019). Lizzeri and Siniscalchi (2008) model the parent as an environmental manipulator who shelters the child from the costs of mistakes, trading off short-term protection against long-term learning. Seror (2022) models the parent as a signal sender and reward provider who communicates the correct action and provides affection when the child's behavior is sufficiently close to the optimum. Cobb-Clark, Salamanca, and Zhu (2019) model parenting style as a household production function with three inputs — time, market goods, and attention — and show that socioeconomic disadvantage constrains the attention-intensive component of parenting by taxing cognitive bandwidth. Heckman and Mosso (2014) identify understanding the mechanisms behind parent-child interactions, and the channels through which parenting influences child development, as crucial tasks for the next generation of studies in human development. This paper formalizes the responsiveness dimension of parenting as a cue-detection problem, identifies four channels through which interventions can improve responsive engagement, and tests the model's predictions in a randomized controlled trial.

Almond, Currie, and Duque (2018) identify the failure to distinguish parental preferences from information constraints as a central gap in the literature on early childhood investment. A substantial empirical literature has established that parental beliefs about the returns to early investment are measurable (Cunha, Elo, and Culhane 2013; Boneva and Rauh 2018; Giannola 2023) and predict parental investment decisions in unitary (Attanasio, Cunha, and Jervis 2019; List, Pernaudet, and Suskind 2021; Conti, Giannola, and Toppeta 2022; Bhalotra, Delavande, Font-Gilabert, and Maselko 2025) or household models (Walker, Grosjean, Cristia, and Delavande 2024). An important open question remains. Can interventions durably change beliefs? Some programs produce belief effects (Cunha, Gerdes, Hu, and Nihtianova 2024; Dupas, Falezan, Jayachandran, and Walsh 2025); others produce behavioral change without belief change (Attanasio, Cunha, and Jervis 2019). The literature lacks a theory of how beliefs are formed, why they predict investments, and why some interventions change beliefs while others do not. This paper provides a theory and tests it: using a randomized controlled trial, we show that an intervention that activates the belief channel, the cue-detection expertise channel, and the trap compression channel moves parental beliefs by 87% of a standard deviation and parental investments by 30% of a standard deviation.

Attanasio (2026) surveys a large literature on the economics of early childhood development. This comprehensive survey discusses parent-directed programs designed to improve the early childhood environment, including home visitation programs, group-based parenting interventions, and psychosocial stimulation programs, such as Reach Up and Learn (Grantham-McGregor, Powell, and Himes 1991; Gertler et al. 2014; Gertler et al. 2021; Attanasio et al. 2014) and the Nurse-Family Partnership (Olds, Henderson, Tatelbaum, and Chamberlin 1986; Heckman, Holland, Makino, Pinto, and Rosales-Rueda 2017). This literature documents substantial heterogeneity in program effectiveness but, as Attanasio (2026) notes, lacks a framework that explains what works for whom and why, one that offers a tighter link between parent-directed programs, theory, econometric methods, and data. The model developed in this paper provides a step in that direction. Any parent-directed program can be characterized by which of the four channels it activates — belief, cue-detection expertise, bandwidth, or trap compression — and the model generates testable predictions about what the program will and will not achieve before an RCT is run. This makes the framework useful not only for researchers evaluating existing programs, but for policymakers designing new ones.

We test the model in the language domain because it is currently the only domain in which available measurement technology distinguishes the cue-contingent component of parental investment from its level at scale. Observational tools such as the Parenting Interactions with Children: Checklist of Observations Linked to Outcomes (PICCOLO; Roggman et al. 2013) and survey-based instruments such as the StimQ Toddler (Cates et al. 2023) capture the responsiveness dimension across domains, but neither is objective nor scalable: PICCOLO requires trained observers and a structured play session, and StimQ relies on caregiver self-report, making both impractical for large-scale naturalistic measurement in school district settings. The LENA system's Advanced Data Extractor records conversational turns decomposed by initiator in the home, continuously, without a research team present, separating child-initiated exchanges, which capture the behavioral signature of a parent reading and responding to child cues, from adult-initiated speech, which reflects level investment regardless of the child's signal. The framework is general; the limitation is the measurement technology, and the language domain is where that technology currently exists at scale.[1]

We test the model's predictions using a randomized controlled trial of the LENA Start Program, a ten-week group-based parenting intervention delivered by the Alief Independent School District in Houston, Texas. The RCT is an experiment in which the model's four testable implications are put to the data. The program is well-suited to this purpose because it activates multiple channels simultaneously: Education and Reporting operate through the belief channel; Coaching and Book

[1] Automatic detection of child-adult interaction is an active area of research, and more recently developed systems such as Lavechin, Bousbib, Bredin, Dupoux, and Cristia (2020) arguably outperform the LENA system on key segmentation tasks (see also Cristia, Bulgarelli, and Bergelson 2020). Our reliance on LENA is not incidental: the feedback mechanism that is central to the intervention, in which parents receive personalized weekly reports on their child's language environment, depends on the fully automated recording, processing, and report generation pipeline that the LENA system provides. No alternative system is currently deployed at this scale in school district settings.

Reading operate through the cue-detection expertise channel; and Feedback operates through the trap compression channel.[2] The model predicts effects on parental beliefs and responsive engagement concentrated in child-initiated conversational exchanges, and null effects on global preferences and general self-efficacy. All four predictions are confirmed: treatment increases the belief factor score by up to 87% of a standard deviation, conversational turns by about 30% of a standard deviation, and produces no detectable effects on global preferences or self-efficacy.

The remainder of the paper is organized as follows. Section 2 develops the model. Section 3 solves it by backward induction and characterizes the self-sealing trap. Section 4 maps the LENA Start Program onto the model's channels and derives four testable implications. Section 5 describes the experimental design, measurement strategy, and data. Section 6 reports the results. Section 7 discusses implications for the beliefs literature and the design of parent-directed programs.

2 Model

A parent invests in the development of a young child through responsive interaction. Responsive parenting involves detecting and responding to child-initiated communicative cues, such as vocalizations, gestures, and bids for attention, in ways that sustain the child's engagement. The developmental value of such interactions depends on two quantities that are not directly observable to the parent: the returns to responsive interaction, γ, and the child's cue rate, θ.

The parameter $\gamma \in \mathbb{R}$ measures the marginal developmental value of a unit of responsive engagement. The child's cue rate $\theta \in [0,1]$ is the probability that a given interaction opportunity generates a detectable, response-worthy cue. The parent holds a reference model for γ and rational beliefs about θ, as described below.

The parent is endowed with bandwidth $b > 0$, which indexes the cognitive and attentional capacity available for responsive engagement. Bandwidth constrains the parent's decisions at every stage: it raises the psychological cost of holding accurate beliefs in Stage 1, and it constrains the intensity of cue detection and response in Stage 2. The parameter $\rho \geqslant 0$ captures cue-detection expertise, that is, the quality of context-specific skill in reading and responding to child communicative bids, which augments what the parent can accomplish under bandwidth constraint. Higher ρ reduces the cognitive cost of cue detection for a given level of b.

Two barriers to responsive parenting. The central insight of the model is that responsive parenting requires the parent to successfully navigate two distinct and sequentially linked barriers, each governed by a different psychological mechanism and each capable of independently block-

[2] Delivering personalized weekly feedback is feasible with the LENA system because recording, scoring, and report generation are fully automated: the parent wears the device at home, uploads the recording, and receives a report without a research team present. An observational instrument such as the PICCOLO could, in principle, capture the responsiveness dimension, but the feedback cycle would require bringing the parent and child to the school district each week, setting up a structured play session, scoring the interaction with a trained observer, and then generating and sharing a report. Such a protocol is not feasible given the resources of a school district program.

ing responsive investment. The model has two stages, solved by backward induction, with each barrier corresponding to one stage.

The *first barrier* is **believing**: the parent must be willing to hold accurate beliefs about γ as her operative working belief. A parent whose prior $(\mu_{\gamma,0}, \sigma^2_{\gamma,0})$ indicates that responsive parenting has high returns may nonetheless distort her working belief downward, because holding an accurate belief about high γ is psychologically painful when bandwidth is low and the shortfall between aspiration and constrained capacity is large. Motivated reasoning, not ignorance, sustains the distorted belief. The degree of distortion depends on the tightness of the prior: a parent with a diffuse prior faces a low accuracy cost and has wide scope for motivated belief revision, while a parent with a tight prior is more constrained. This is the mechanism governing Stage 1.

The *second barrier* is **acting**: even a parent who holds accurate prior beliefs about γ may fail to translate them into responsive engagement if her working belief falls below the threshold at which reading child cues becomes worthwhile. Below this threshold, the parent responds uniformly regardless of the child's cues, generating no calibration evidence to sustain or reinforce accurate beliefs. The trap is self-sealing. This is the mechanism governing Stage 2.

These two barriers interact. Low bandwidth simultaneously increases the psychological pain of accurate beliefs and degrades the quality of engagement. The model identifies four channels through which interventions can address the two barriers, each targeting a distinct element of the model and each capable of moving the parent toward responsive engagement independently or in combination. Understanding this structure is essential for designing interventions that work, and for understanding why interventions that activate only one channel may fall short when the other barriers remain intact.

2.1 Stage 1: Motivated Belief Distortion

2.1.1 Reference Prior

The parent holds a reference prior over the returns to responsive interaction γ. We model this as a Normal distribution,

$$\gamma \sim \mathcal{N}(\mu_{\gamma,0}, \sigma^2_{\gamma,0}) \tag{1}$$

where the choice of family is made for tractability and is without loss of generality for the qualitative results.[3] The prior mean $\mu_{\gamma,0} \in \mathbb{R}$ is the parent's baseline belief about the returns to responsive interaction, representing accumulated knowledge from upbringing, cultural norms, social networks, and prior experience with child development.[4] The prior variance $\sigma^2_{\gamma,0} > 0$ captures baseline

[3] As we show in Section 2.1.3, the accuracy cost D_γ coincides with the Kullback-Leibler divergence between two Normal distributions, which justifies the precision-scaled quadratic specification introduced below.

[4] The model takes the reference prior mean $\mu_{\gamma,0}$ as given and treats departures of the working belief, which we introduce below, from $\mu_{\gamma,0}$ as belief distortion. This requires an implicit assumption: that $\mu_{\gamma,0}$ is the relevant benchmark against which distortion is measured. Two interpretations are consistent with this assumption. Under the first, γ varies across families and $\mu_{\gamma,0}$ accurately reflects the family-specific marginal productivity of responsive investment, so that distortion is a departure from the true return for a given family. Under the second, $\mu_{\gamma,0}$ may itself be a biased prior,

uncertainty about γ: a large $\sigma^2_{\gamma,0}$ reflects substantial uncertainty about whether and how much responsive parenting matters, while a small $\sigma^2_{\gamma,0}$ reflects a tight prior. Both parameters vary across parents, reflecting heterogeneity in prior exposure to information about child development.

This parametrization is empirically grounded. A first body of evidence establishes that parental beliefs about the returns to investment are heterogeneous and consequential. Cunha, Elo, and Culhane (2013) first documented systematic heterogeneity in maternal beliefs about both child potential and the returns to parental investment. Attanasio, Cunha, and Jervis (2019) provide the most direct evidence: using elicited subjective production functions from a randomized controlled trial in Colombia, they show that mothers systematically underestimate the marginal productivity of their investments and that these beliefs predict actual investment decisions. A second body of evidence establishes that this heterogeneity follows a clear socioeconomic gradient. List, Pernaudet, and Suskind (2021) show that higher-SES parents are substantially more likely to believe that parental investments affect child development, with low-SES parents' beliefs averaging half a standard deviation below those of higher-SES parents. Boneva and Rauh (2018) show that belief heterogeneity varies by household income and contributes to intergenerational earnings persistence. Most directly relevant to our model, Bhalotra, Delavande, Font-Gilabert, and Maselko (2025) show that the SES gap in parenting investment is driven by heterogeneity in *beliefs* about developmental returns rather than by heterogeneity in *preferences* for child outcomes—precisely the mechanism through which $\mu_{\gamma,0}$ generates investment differences in our framework.

2.1.2 Psychological Cost of Accurate Beliefs

The parent enters Stage 1 facing a decision that is absent from standard models: she must choose the working belief $\hat{\mu}_\gamma$ that will govern her behavior in Stage 2, knowing that accurate beliefs may be psychologically costly to hold. We model this cost through a shortfall between what the parent believes she could achieve and what her constraints allow her to deliver.

To define this shortfall, let $V_2(\hat{\mu}_\gamma, b, \rho)$ denote the Stage 2 continuation value, which integrates over the child's cue rate θ through the information structure introduced in Section 2.2. Define $V_2^*(\hat{\mu}_\gamma)$ as the value the parent would achieve if cognitive resources were free, that is, if the attention cost were zero.[5] When the parent operates under a bandwidth constraint, V_2 falls short of V_2^*. A parent who believes responsive parenting matters but knows she cannot fully deliver it confronts a painful discrepancy between aspiration and capacity. We define the psychological cost

for example, one that underestimates γ due to limited exposure to information about child development. The model's predictions are identical under both interpretations, since what drives these predictions is the relationship between $\mu_{\gamma,0}$ and the threshold for the corner regime we derive below, not whether $\mu_{\gamma,0}$ equals the true γ. The paper uses the language of accurate and distorted beliefs as shorthand for the relationship between prior and working beliefs without taking a stand on whether $\mu_{\gamma,0}$ itself is unbiased.

[5]Under this benchmark, uncertainty about θ can be resolved through the signal at zero cost. The benchmark is endogenous, derived from the model's primitives rather than imposed as an exogenous aspiration; its closed-form expression is established in Section 3.1.1. Crucially, we show that V_2^* does not depend on b or ρ, since these parameters enter only through the attention cost.

as a scaled version of this shortfall:

$$G(\hat{\mu}_\gamma, b, \rho) \;\equiv\; \psi \cdot \left[V_2^*(\hat{\mu}_\gamma) \;-\; V_2(\hat{\mu}_\gamma, b, \rho)\right] \tag{2}$$

where $\psi > 0$ governs the intensity of the psychological response. As we establish in Section 3.2.1, G is increasing in $\hat{\mu}_\gamma$: a parent with higher beliefs about the returns to responsive interaction perceives a larger shortfall between what she aspires to achieve and what her constraints permit. This is the motive for downward belief distortion, not a preference for false beliefs per se, but a rational response to the aversiveness of accurate beliefs that cannot be acted upon.

Three bodies of evidence support this structure. First, a mature theoretical literature establishes that agents trade off the hedonic benefits of distorted beliefs against the decision costs of inaccuracy (Bénabou and Tirole 2002; Bénabou 2015; Brunnermeier and Parker 2005; Kőszegi 2006). In these models, belief distortion is not a cognitive failure but a rational response to the psychological costs of accurate self-assessment.

Second, experimental evidence confirms that accurate beliefs are psychologically costly and that individuals distort beliefs in response. Celik-Katreniak (2021) show that improving self-assessment accuracy is associated with lower happiness even when it improves performance, direct evidence for the psychological cost of accurate beliefs that our model formalizes as G. Castagnetti and Schmacker (2022) show that belief distortion is specific to ego-relevant domains and absent in control conditions, consistent with the model's assumption that G operates through domain-specific aspiration-capacity gaps rather than general confidence. Drobner and Goerg (2024) show that individuals who receive poor feedback on an IQ test retroactively downplay the importance of the test to protect their self-image, and update beliefs more optimistically than Bayesian benchmarks when the task is framed as highly relevant to their self-worth. Thunström, Nordström, Shogren, Ehmke, and Veld (2016) extend this to the limiting case: individuals choose to remain ignorant of accurate self-threatening information even when acquiring it is costless, consistent with a model in which accurate beliefs are avoided when acting on them is infeasible.

Third, evidence from aspiration research provides direct empirical grounding for the shortfall structure captured by the G function we introduce below. Lim (2020) shows that chronic illness widens the gap between current life evaluation and desired future outcomes, generating psychological distress that is increasing in the gap and more severe among low-income individuals. This finding suggests a structure that maps directly onto G, which is increasing in the shortfall and larger when bandwidth b is smaller. Hanell (2022) provides complementary evidence using data from over 14,000 individuals across ten western European countries: when life outcomes fall short of self-enhancement aspirations, specifically aspirations for success and wealth, the gap itself reduces life satisfaction above and beyond objective circumstances, consistent with a psychological cost that is driven by the shortfall rather than by the level of deprivation alone. We now formalize this structure.

2.1.3 Accuracy Cost of Belief Distortion

Departing from the reference prior is costly. The parent bears an accuracy cost that is increasing in the distance between her working belief $\hat{\mu}_\gamma$ and her reference prior mean $\mu_{\gamma,0}$. We specify this as a quadratic penalty scaled by prior precision:

$$D_\gamma(\hat{\mu}_\gamma, \mu_{\gamma,0}, \sigma^2_{\gamma,0}) \;=\; \frac{\lambda_\gamma}{2\sigma^2_{\gamma,0}}(\hat{\mu}_\gamma - \mu_{\gamma,0})^2 \tag{3}$$

where $\lambda_\gamma > 0$ scales the overall cost. This specification coincides exactly with the Kullback-Leibler divergence from the working belief to the reference prior under the Normal family assumed in equation (1), so the precision-scaled quadratic is not an arbitrary choice but a direct consequence of the distributional assumption.[6] The precision scaling $1/\sigma^2_{\gamma,0}$ is the key feature of D_γ. A parent with a tight prior faces a high cost of sustaining a distorted working belief and is therefore constrained in how far she can shade $\hat{\mu}_\gamma$ away from $\mu_{\gamma,0}$. A parent with a diffuse prior faces a low accuracy cost and has wide scope for motivated belief revision. This is the channel through which heterogeneity in prior information shapes the degree of belief distortion: parents who are better informed about the returns to responsive interaction, those with smaller $\sigma^2_{\gamma,0}$, distort less. This delivers a testable prediction: information interventions that tighten parental priors should reduce the scope for motivated reasoning, not merely shift its center.

2.1.4 Belief Distortion Problem

The parent chooses $\hat{\mu}_\gamma$ to maximize an objective that trades off three forces: the continuation value of Stage 2 engagement, the psychological cost of accurate beliefs, and the accuracy cost of departing from the reference prior:

$$\max_{\hat{\mu}_\gamma} \quad V_2(\hat{\mu}_\gamma, b, \rho) \;-\; \frac{\lambda_\gamma}{2\sigma^2_{\gamma,0}}(\hat{\mu}_\gamma - \mu_{\gamma,0})^2 \;-\; G(\hat{\mu}_\gamma, b, \rho) \tag{4}$$

The structure of problem (20) reveals the central tension of Stage 1. Higher $\hat{\mu}_\gamma$ raises V_2 by inducing greater engagement in Stage 2, but also raises G by widening the perceived shortfall between aspiration and capacity. The accuracy cost D_γ disciplines the downward adjustment: a tight prior makes belief distortion expensive, while a diffuse prior leaves wide scope for motivated revision. The optimal working belief $\hat{\mu}^*_\gamma$ and its comparative statics are characterized in Section 3.2, where we show in particular that bandwidth-constrained parents distort more severely and that the degree of distortion is decreasing in prior precision.

[6] Hyland and Albarracin (2025) develop a resource-rational framework in which agents weigh belief utility against the KL-divergence cost of revising their prior, providing independent theoretical support for this specification.

2.2 Stage 2: Cue Detection and Responsive Investment

In Stage 2, the parent observes opportunities to interact with the child and decides how attentively to monitor for response-worthy cues. The working belief $\hat{\mu}_\gamma$ chosen in Stage 1 enters Stage 2 as a parameter: the parent takes her operative belief about the returns to responsive interaction as given and optimizes engagement intensity conditional on it.

2.2.1 Payoff Function

The primitive payoff given response intensity $x > 0$ and cue intensity θ is:

$$u(x,\theta) = \hat{\mu}_\gamma \, \theta \, \ln x - \frac{1}{2\delta} \left(\ln x - \ln x^R\right)^2 \tag{5}$$

The first term is the developmental return: the product of the believed return to responsiveness ($\hat{\mu}_\gamma$), the cue intensity (θ), and the log of response intensity ($\ln x$). The inclusion of this term in the payoff function is grounded in an extensive developmental literature establishing that it is the contingent, cue-calibrated nature of parental responses to child-initiated bids, not the quantity of investment alone, that drives early language and cognitive development. Romeo et al. (2018) demonstrate that conversational turns predict activation in the left inferior frontal gyrus (Broca's area, a brain region critical for language processing) during story listening, independently of socioeconomic status, IQ, and total adult speech, and that a family-based intervention increasing conversational turns produced structural brain changes in regions associated with language development. Luchkina and Xu (2024) and McGillion et al. (2013) show that only utterances that are both semantically appropriate and temporally contingent on infant vocalizations predict vocabulary growth, while the total amount or simplicity of language input does not. Tamis-LeMonda, Kuchirko, and Song (2014) explain the mechanism: temporal contiguity and contingency facilitate word-to-referent mapping by demonstrating the communicative nature of language, not merely exposing the child to more words. The payoff function (5) captures this directly. It is zero if the child emits no cue ($\theta = 0$), regardless of the parent's response intensity, and it is maximized when the parent's response is calibrated to the child's actual cue rate.

The distance function $\frac{1}{2\delta}(\ln x - \ln x^R)^2$ captures the cost of departing from the parent's habitual mode of engagement. This cost is distinct from the bandwidth cost we introduce below. The distance function captures the cost of operating at a responsiveness level that differs from the parent's established behavioral default. Three complementary bodies of evidence motivate this assumption.

First, parenting behavior is highly routinized, and departures from established routines are costly. Wood, Tam, and Guerrero Witt (2005) show that changing circumstances disrupts learned response patterns and shifts behavior toward intentional control, implying that operating outside a habitual mode requires deliberate effort that is not required within it. Thomas, Poortinga, and Sautkina (2016) provide direct evidence that habits weaken under sufficiently large contextual disruptions but reassert themselves as the new environment becomes stable, consistent with automatic processes taking over and reducing the influence of conscious intentions. This dynamic supports modeling the habitual default as a strong attractor from which large departures are disproportionately costly to sustain. Chetty and Szeidl (2016) formalize this as a consumption commitments model in which adjustment costs generate smooth dynamics for small shocks but qualitatively different responses for large ones. Their structure also motivates the convex penalty in $(\ln x - \ln x^R)^2$.

Second, responding more contingently than the habitual default requires overriding automatic parenting responses, and this override is effortful independently of attention allocation. Zhang, Gatzke-Kopp, Cole, and Ram (2022) conceptualize effective parenting in the context of challenging child behavior as requiring the inhibition of prepotent reactions through higher-order executive processes, allowing parents to balance the external demand of responding to the child with the internal demand of restoring their own physiological equilibrium, thus explicitly distinguishing this self-regulatory demand from attentional constraints. Tate, Trofholz, Youngblood, Goldschmidt, and Berge (2024) provide ecological momentary assessment evidence that resource depletion earlier in the day predicts more coercive parenting practices at dinner, consistent with depleted self-regulatory capacity pushing parents toward reactive defaults.[7] The cost is symmetric: just as deviating below x^R involves reverting to reactive defaults, deviating above x^R involves sustaining a level of contingent responsiveness that exceeds what is automatic, which is effortful in the same way.

Third, departures from one's own habitual engagement standard generate psychic costs through self-evaluative comparison. Stone and Cooper (2001) show that dissonance arises when people evaluate behavior against personal idiosyncratic standards represented in memory, explicitly distinguishing this self-anchored comparison from deviation from external norms. Oh, Lee, and Cho (2023) provide direct evidence that larger discrepancies between actual and ideal self-concepts generate greater shame, while actual/ought discrepancies generate greater guilt, consistent with psychic costs that are increasing in the magnitude of deviation from internalized self-standards. In parenting contexts specifically, Liss, Schiffrin, and Rizzo (2013) and Kenny and Robinson (2026) show that maternal self-discrepancies predict guilt and shame, and Slobodin, Cohen, Arden, and Katz (2020) shows that maternal guilt acts as a self-evaluative reaction associated with an increased use of controlling parenting practices.

The three components are complementary: habit persistence makes the default easy to execute and costly to depart from, self-regulation costs make overriding automatic responses effortful, and psychic inconsistency costs generate negative affect when behavior departs from internalized self-standards. Together they justify a departure cost that is smooth, symmetric, and convex in

[7]The ego depletion literature on which this argument draws has faced replication challenges (Hagger et al. 2016). Our assumption rests on the weaker claim that overriding automatic responses requires effortful control, which is supported directly by parenting-specific evidence (Zhang, Gatzke-Kopp, Cole, and Ram 2022; Tate, Trofholz, Youngblood, Goldschmidt, and Berge 2024; Geeraerts et al. 2021) independently of the broader ego depletion hypothesis.

$(\ln x - \ln x^R)$, with the log-quadratic form $\frac{1}{2\delta}(\ln x - \ln x^R)^2$ adopted as a tractable representation of these properties.

Parenting interventions can raise x^R through two distinct mechanisms. The first is *repertoire building*: programs that teach parents to integrate interaction activities into daily routines (narrating during cooking, naming objects during shopping, dialogic reading at bedtime) expand the parent's default repertoire of engagement, raising her habitual baseline regardless of the child's specific cue (Grantham-McGregor, Powell, and Himes 1991). The second is *target anchoring*: programs that set specific, quantitative goals for parental engagement (as it is done in the intervention we study in this paper) anchor the parent's reference level to a higher benchmark. Both mechanisms raise x^R, improving the parent's default level of engagement regardless of whether the interaction is calibrated to the child's specific cue.

2.2.2 Cue Detection and Information Structure

The parent's prior over θ is:

$$\theta \sim \text{Beta}(\alpha, \beta), \tag{6}$$

where α and β are determined by the rational posteriors μ_θ and σ_θ^2 via the standard Beta-distribution moment conditions:

$$\mu_\theta = \frac{\alpha}{\alpha+\beta}, \qquad \sigma_\theta^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

In a given interaction episode of length n, the number of cues s the parent detects is distributed as:

$$s \mid \theta \sim \text{Binomial}(n, \theta). \tag{7}$$

Integrating over the Beta prior, the marginal distribution of s is Beta-Binomial. The episode length n is the parent's choice variable: it indexes the attentional investment devoted to monitoring for cues, and therefore the precision with which the parent reads the child's signal.

The outcome of the parent's attention choice is a posterior belief about the child's cue rate θ. After observing s cues in an episode of length n, the parent updates her prior $\theta \sim \text{Beta}(\alpha, \beta)$ using Bayes' rule. Under the Beta-Binomial structure, the posterior distribution of θ given s remains Beta with updated parameters, and the posterior mean is:

$$\mathbb{E}(\theta \mid s) = \omega_\theta \mu_\theta + (1 - \omega_\theta)\frac{s}{n} \tag{8}$$

It is this posterior mean that enters the parent's optimal response in Stage 2, linking the attention choice ω_θ to the behavioral outcome $x^*(s)$.

The Beta distribution is the conjugate prior for the Binomial likelihood, so the parent's posterior after observing s cues in n opportunities remains Beta, with updated parameters that incorporate the new evidence. This conjugate structure delivers a simple relationship between the episode length n and the posterior-to-prior variance ratio ω_θ:

$$\omega_\theta = \frac{\alpha + \beta}{\alpha + \beta + n} \tag{9}$$

More monitoring — a longer episode n — reduces ω_θ, tightening the posterior around the observed cue frequency. No monitoring — $n = 0$ — leaves $\omega_\theta = 1$ and the posterior equal to the prior.

This conjugate structure is not only an assumption imposed for tractability, but it is also what standard Bayesian updating delivers when the unknown parameter is a probability and the data are counts. Equation (9) establishes a one-to-one, monotone decreasing relationship between the episode length n and the posterior-to-prior variance ratio ω_θ: every value of $n \geqslant 0$ maps to a unique $\omega_\theta \in (0, 1]$, and every $\omega_\theta \in (0, 1]$ corresponds to a unique $n \geqslant 0$. The parent's problem can therefore be equivalently stated as a choice over ω_θ rather than n, which is the formulation adopted in Section 3. This reparameterization has two advantages: it directly connects the parent's attention choice to the precision of her posterior belief about θ, and it allows the cost function $C(\omega_\theta) = -\ln \omega_\theta$ to be expressed in terms of the informativeness of the signal rather than the raw episode length.

Finally, the Beta-Binomial structure treats the child's cue rate θ as an exogenous property of the child, abstracting from models in which the child is a strategic agent who responds to parental behavior (Seror 2022; Lizzeri and Siniscalchi 2008). For children about two years of age, which is the population we study, this abstraction is a reasonable approximation: communicative bids at this developmental stage are arguably driven by the child's internal state and developmental level rather than by strategic responses to parental engagement.

2.2.3 Attention Cost Function

The unit cost of information is:

$$\kappa(b, \rho) = \frac{\kappa_0}{b(1 + \rho)} \tag{10}$$

where $\kappa_0 > 0$ is the base cost of cognitive processing, $b > 0$ is the parent's bandwidth, and $\rho \geqslant 0$ is the parent's cue-detection expertise in the current context. Bandwidth b captures the parent's general cognitive resources available for parenting at any given moment. It is depleted by competing demands of economic survival, including financial stress, irregular work schedules, household management, fatigue, and the cognitive load of navigating poverty. Shah, Mullainathan, and Shafir (2012) document that scarcity taxes cognitive bandwidth, reducing performance across domains that require deliberate attention. Mani, Mullainathan, Shafir, and Zhao (2013) provide experimental evidence that financial concerns directly impair cognitive function, reducing performance on fluid intelligence and impulse control tasks. In our model, low b means that every unit of information the parent acquires about the child's cue is expensive, not because the child's signals are hard to read in an absolute sense, but because the parent's cognitive resources are consumed by other demands. A parent who is worried about rent, managing multiple children, and cooking dinner simultaneously has little bandwidth left for reading a toddler's subtle cues. Cheremukhin, Popova, and Tutino (2011) document substantial heterogeneity in individual-level attention costs, consistent with our model's prediction that b varies across parents.

Cue-detection expertise ρ captures the parent's capacity to notice, interpret, and respond to child communicative bids, built through deliberate practice of serve-and-return interactions. A parent with high ρ has developed perceptual sensitivity to her child's signals: she notices when the child shifts attention, waits for the child to initiate before responding, and recognizes when one interaction ends and another begins. These skills reduce the effective attention cost because cue detection has become partly automatic — the parent no longer needs to deploy deliberate cognitive effort to read signals that once required sustained, bandwidth-consuming attention. The Filming Interactions to Nurture Development (FIND) program provides a direct empirical instantiation of ρ-raising: through video coaching, parents practice noticing and shifting attention toward the child's cue, waiting for the child's serve before responding, and sustaining back-and-forth exchanges (Fisher, Frenkel, Noll, Berry, and Yockelson 2016). By showing parents objective video evidence of their own successful serve-and-return interactions, FIND builds automated cue-detection patterns — the context-specific expertise that ρ captures — and is anticipated to increase caregivers' perceptual sensitivity to infant emotional cues.

The multiplicative structure $b(1+\rho)$ implies that bandwidth and cue-detection expertise jointly determine the effective cognitive resources available for cue detection. The $1+\rho$ specification ensures that the cost is finite when $\rho = 0$ (no expertise): the parent can still read cues, just at the full bandwidth cost κ_0/b. When $\rho > 0$, each unit of cue-detection expertise reduces the effective cost proportionally. A parent with high ρ in a given context behaves as if she had more bandwidth, not because the competing demands on her attention have diminished, but because cue detection has been automated and no longer competes for the same cognitive resources.

2.2.4 Cost of Monitoring

Sustained attentive monitoring is costly. We parameterize the cost of choosing episode length n in terms of the implied posterior precision. Let $\omega_\theta \in (0, 1]$ denote the ratio of posterior to prior variance of θ after observing s cues in n opportunities. The cost of monitoring is:

$$C(\omega_\theta) = -\ln \omega_\theta \geqslant 0. \tag{11}$$

This log precision ratio cost is increasing in monitoring intensity (ω_θ decreasing) and equals zero when $\omega_\theta = 1$, that is, when the parent exerts no monitoring effort and the posterior equals the prior.

The cost function (11) is grounded in the rational inattention framework introduced by Sims

(2003), in which agents face Shannon mutual information costs when processing signals about the state of the world. The log precision ratio cost $C(\omega_\theta) = -\ln\omega_\theta$ is a direct implementation of this framework: it is proportional to the mutual information $I(\theta; s) = -\frac{1}{2}\ln\omega_\theta \geqslant 0$ between the child's cue rate θ and the signal s, with $C(\omega_\theta) = 2I(\theta; s)$, scaled by the unit cost $\kappa(b,\rho)$. Caplin, Dean, and Leahy (2022) provide axiomatic foundations showing that Shannon entropy is pinpointed by two behavioral axioms (Locally Invariant Posteriors and Only Payoffs Matter) establishing the theoretical uniqueness of this cost function among the class of posterior-separable costs.

2.2.5 Information Acquisition and Optimal Responsive Investment

The child emits a cue of intensity $\theta \in [0,1]$, which the parent does not observe directly. The parent first chooses how much attention to allocate to reading the cue (the exposure n), then observes a signal s, and then chooses how intensely to respond (x). The inner expectation $\mathbb{E}[\cdot \mid s]$ integrates over θ given the signal s, and $\mathbb{E}_s$ then integrates over signal realizations. The Stage 2 value function therefore integrates over all realizations of θ by construction, which is consistent with its role as a continuation value in the Stage 1 problem of Section 2.1. The parent's full problem is:

$$V_2(\hat{\mu}_\gamma, b, \rho) = \max_{n \geqslant 0} \; \mathbb{E}_s\left[\max_{x>0} \; \mathbb{E}\left[\hat{\mu}_\gamma\,\theta\,\ln x - \frac{1}{2\delta}\left(\ln x - \ln x^R\right)^2 \;\Big|\; s\right]\right] - \kappa(b,\rho)\cdot C(n) \qquad (12)$$

subject to (6), (7), (10), and (11).

3 Solution

We solve the model by backward induction, beginning with Stage 2 (Section 3.1) and proceeding to Stage 1 (Section 3.2). The backward induction solution reveals that low bandwidth simultaneously facilitates belief distortion and prevents cue detection. These two barriers interact through recursive complementarity to produce a self-sealing trap, characterized in Section 3.3. The model identifies four channels through which interventions can address the trap: the **belief channel**, which raises the parent's prior mean $\mu_{\gamma,0}$ above the critical prior mean $\bar{\mu}_{\gamma,0}$; the **bandwidth channel**, which raises cognitive bandwidth b, lowering both $\bar{\mu}_{\gamma,0}$ and the psychological cost of accurate beliefs; the **cue-detection expertise channel**, which raises cue-detection expertise ρ, with the same dual effect as the bandwidth channel; and the **trap compression channel**, which raises the reference investment $\ln x^R$, lowering $\bar{\mu}_{\gamma,0}$ without affecting the psychological cost function. Each channel is targetable by a different type of intervention.

3.1 Stage 2: Cue Detection and Responsive Investment

Define the attention cost threshold:

$$\bar{\mu}_\gamma \equiv \sqrt{\frac{2\kappa(b,\rho)}{\delta\sigma_\theta^2}} = \sqrt{\frac{2\kappa_0}{\delta\sigma_\theta^2\, b(1+\rho)}} \tag{13}$$

We begin with a lemma that characterizes the key statistical property of the Beta-Binomial signal structure.

Lemma 1 (Variance of the Beta-Binomial posterior mean)**.** *Under the Beta-Binomial structure with prior* $\theta \sim \text{Beta}(\alpha, \beta)$ *and* $s \mid \theta \sim \text{Binomial}(n, \theta)$*, the variance of the posterior mean* $\mathbb{E}(\theta \mid s)$ *over signal realizations* s *is:*

$$\text{Var}_s[\mathbb{E}(\theta \mid s)] = \sigma_\theta^2 \cdot \frac{n}{\alpha+\beta+n} = \sigma_\theta^2(1-\omega_\theta) \tag{14}$$

where $\sigma_\theta^2 = \alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$ *is the prior variance and* $\omega_\theta = (\alpha+\beta)/(\alpha+\beta+n) \in (0,1]$ *is the ratio of posterior to prior variance.*

Proof. See Appendix A. □

Proposition 1 (Stage 2 optimal choices)**.** *Given working belief* $\hat{\mu}_\gamma$*, bandwidth* b*, and cue-detection expertise* ρ*, the parent's optimal monitoring precision and optimal response are:*

(i) ***Optimal monitoring precision.***

$$\omega_\theta^* = \begin{cases} \dfrac{2\kappa(b,\rho)}{\delta\hat{\mu}_\gamma^2\,\sigma_\theta^2}, & \textit{if} \quad \hat{\mu}_\gamma > \bar{\mu}_\gamma \\ 1, & \textit{otherwise} \end{cases} \tag{15}$$

(ii) ***Optimal response conditional on signal.***

$$\ln x^*(s) = \begin{cases} \ln x^R + \delta\hat{\mu}_\gamma \left[\dfrac{2\kappa(b,\rho)}{\delta\hat{\mu}_\gamma^2\,\sigma_\theta^2}\,\mu_\theta + \left(1 - \dfrac{2\kappa(b,\rho)}{\delta\hat{\mu}_\gamma^2\,\sigma_\theta^2}\right)\dfrac{s}{n} \right], & \textit{if} \quad \hat{\mu}_\gamma > \bar{\mu}_\gamma \\ \ln x^R + \delta\hat{\mu}_\gamma\,\mu_\theta, & \textit{otherwise} \end{cases} \tag{16}$$

In the interior regime, the optimal response is a weighted average of a uniform response based on the prior mean μ_θ *and a signal-calibrated response based on the observed frequency* s/n*, with the weight on the signal increasing as* $\hat{\mu}_\gamma$ *rises above* $\bar{\mu}_\gamma$*. In the corner regime, the parent responds uniformly regardless of the signal.*

Proof. See Appendix A. □

The corner solution $\omega_\theta^* = 1$ — in which the parent acquires no information about child cues and responds uniformly — is a natural prediction of the rational inattention framework. Mackowiak, Matejka, and Wiederholt (2018) show that when the stakes associated with a particular dimension of the state are sufficiently low relative to the cost of processing, agents optimally choose zero precision on that dimension, responding uniformly. Matějka and McKay (2015) establish that with Shannon costs and discrete actions, rationally inattentive agents may assign exactly zero attention to some alternatives, producing behavior in which some states are effectively ignored. In our model, the parent's working belief $\hat{\mu}^*_\gamma$ falling below a threshold, which we derive below, is the condition under which the value of information about θ falls below the processing cost $\kappa(b,\rho)$, making the corner solution optimal.

Proposition 2 (Comparative statics of ω^*_θ)**.** *In the interior regime ($\hat{\mu}_\gamma > \bar{\mu}_\gamma$):*

(i) $\partial\omega^*_\theta/\partial\hat{\mu}_\gamma < 0$*, with the effect proportional to* $\hat{\mu}_\gamma^{-3}$*, so it is strongest at low beliefs.*

(ii) $\partial\omega^*_\theta/\partial b < 0$*, with the effect strongest when bandwidth is most scarce.*

(iii) $\partial\omega^*_\theta/\partial\rho < 0$*, with the effect strongest when cue-detection expertise is lowest.*

Proof. See Appendix A. □

3.1.1 Indirect Utility

Substituting ω^*_θ from Proposition 1 into the Stage 2 objective, the parent's indirect utility is:

$$V_2(\hat{\mu}_\gamma, b, \rho) = \begin{cases} \hat{\mu}_\gamma\,\mu_\theta\,\ln x^R + \frac{\delta}{2}\hat{\mu}_\gamma^2\,\mu_\theta^2 + \frac{1}{2}\left(\delta\hat{\mu}_\gamma^2\sigma_\theta^2 - 2\kappa(b,\rho)\right) + \kappa(b,\rho)\ln\frac{2\kappa(b,\rho)}{\delta\hat{\mu}_\gamma^2\sigma_\theta^2}, & \text{if} \quad \hat{\mu}_\gamma > \bar{\mu}_\gamma \\ \hat{\mu}_\gamma\,\mu_\theta\,\ln x^R + \frac{\delta}{2}\hat{\mu}_\gamma^2\,\mu_\theta^2, & \text{otherwise} \end{cases} \tag{17}$$

The function V_2 is C^1, but not C^2 in $\hat{\mu}_\gamma$ (see Proposition B1, Appendix B). The second derivative jumps by $2\delta\sigma_\theta^2$ at the threshold $\bar{\mu}_\gamma$, reflecting the discrete activation of the calibration channel as beliefs cross the threshold. In the interior regime, believed returns are complementary to both bandwidth and cue-detection expertise, see Proposition B2, Appendix B).

The unconstrained benchmark $V_2^*(\hat{\mu}_\gamma)$, promised in Section 2.1, is obtained by evaluating V_2 as $\kappa(b,\rho) \to 0$:

$$V_2^*(\hat{\mu}_\gamma) = \hat{\mu}_\gamma\,\mu_\theta\,\ln x^R + \frac{\delta}{2}\hat{\mu}_\gamma^2\,\mu_\theta^2 + \frac{\delta}{2}\hat{\mu}_\gamma^2\,\sigma_\theta^2 \tag{18}$$

Note that V_2^* does not depend on b or ρ, consistent with the discussion in Section 2.1.

3.1.2 Discussion

The threshold $\bar{\mu}_\gamma$ is the critical belief at which the parent transitions from not reading cues to reading cues. Proposition B1 establishes that this transition is accompanied by a discrete jump in the curvature of V_2 with respect to $\hat{\mu}_\gamma$. Below $\bar{\mu}_\gamma$, the curvature equals $\delta\mu_\theta^2$, reflecting only the level channel: additional believed returns raise the intensity of response but not its sensitivity to child cues. Above $\bar{\mu}_\gamma$, the curvature jumps to $\delta(\mu_\theta^2 + \sigma_\theta^2(1+\omega_\theta^*))$, reflecting both the level channel and the calibration channel. Since $\omega_\theta^* \in (0,1]$ in the interior, the calibration channel adds between $\delta\sigma_\theta^2$ and $2\delta\sigma_\theta^2$ to the curvature.

This curvature jump has a direct implication for Stage 1. The continuation value V_2 enters the Stage 1 tradeoff between the accuracy cost and the psychological cost of the shortfall. Below $\bar{\mu}_\gamma$, the material value of beliefs grows slowly, and the psychological motive to distort beliefs downward may dominate. Above $\bar{\mu}_\gamma$, the material value grows more rapidly and the accuracy motive strengthens. This creates a discontinuity in the incentive to distort: a parent just above the threshold has more to lose from distorting downward than one just below, even though the two are close in belief space.

Proposition B2 characterizes the joint productivity of the parameters that govern responsive parenting. Parts (i) and (ii) establish that believed returns are complementary to both bandwidth and cue-detection expertise: a parent who believes responsiveness matters gains more from additional expertise, and a parent with greater cue-detection expertise gains more from accurate beliefs. Part (iii) identifies a regime dependence in the complementarity between bandwidth and cue-detection expertise: the two are complements for constrained parents near the corner ($\omega_\theta^* > 1/e$) and substitutes for well-resourced parents deep in the interior ($\omega_\theta^* < 1/e$). For parents operating near the corner, all inputs to responsive parenting are pairwise complements, implying that interventions targeting such parents should treat their components as mutually reinforcing rather than substitutable.

The threshold $\bar{\mu}_\gamma$ clarifies a general principle for intervention design. A program that raises $\hat{\mu}_\gamma$ alone increases the parent's believed returns but leaves $\bar{\mu}_\gamma$ unchanged; if the initial $\hat{\mu}_\gamma$ is sufficiently below $\bar{\mu}_\gamma$, the threshold is never crossed and cue detection is unaffected. Conversely, a program that raises b or ρ alone lowers $\bar{\mu}_\gamma$ but does not raise $\hat{\mu}_\gamma$. Either component, acting in isolation, may therefore fail to move the parent across the threshold if the gap is too large. An intervention that simultaneously raises $\hat{\mu}_\gamma$ and lowers $\bar{\mu}_\gamma$ addresses both barriers at once, and the complementarities in Proposition B2 imply that such components are mutually reinforcing.

3.2 Stage 1: Motivated Belief Distortion

3.2.1 Psychological Cost Function

We begin by deriving the closed form of the psychological cost G introduced in Section 2.1.2. Substituting the expressions for V_2^* and V_2 from equations (18) and (17) into $G(\hat{\mu}_\gamma, b, \rho) = \psi[V_2^*(\hat{\mu}_\gamma) - V_2(\hat{\mu}_\gamma, b, \rho)]$:

$$G(\hat{\mu}_\gamma, b, \rho) = \begin{cases} \psi\,\kappa(b,\rho)\left(1 + \ln \dfrac{\delta \hat{\mu}_\gamma^2 \sigma_\theta^2}{2\kappa(b,\rho)}\right), & \text{if} \quad \hat{\mu}_\gamma > \bar{\mu}_\gamma \\ \dfrac{\psi\delta}{2}\,\hat{\mu}_\gamma^2 \sigma_\theta^2, & \text{otherwise} \end{cases} \tag{19}$$

In the interior regime, G reflects the calibration loss from imperfect cue detection, scaled by ψ. In the corner regime, the parent forgoes the entire calibration gain $\frac{\delta}{2}\hat{\mu}_\gamma^2\sigma_\theta^2$, so G is larger. Note that in the corner regime G is independent of b and ρ, a property that will play a central role in the characterization of distortion below.

Proposition 3 (Properties of G)**.** *The psychological cost function* $G(\hat{\mu}_\gamma, b, \rho)$ *has the following properties:*

(i) ***Continuity and differentiability.*** G *is continuous and differentiable at the threshold* $\hat{\mu}_\gamma = \bar{\mu}_\gamma$*.*

(ii) ***Increasing in*** $\hat{\mu}_\gamma$***.*** $\partial G/\partial\hat{\mu}_\gamma > 0$ *in both regimes: the parent can reduce the psychological cost by distorting* $\hat{\mu}_\gamma$ *downward.*

(iii) ***Decreasing in*** b ***and*** ρ***.*** G *is decreasing in bandwidth* b *and cue-detection expertise* ρ *in the interior regime, through* $\kappa(b,\rho) = \kappa_0/[b(1+\rho)]$*. In the corner regime,* $G = \frac{\psi\delta}{2}\hat{\mu}_\gamma^2\sigma_\theta^2$ *is independent of* b *and* ρ*.*

(iv) ***Larger shortfall in the corner.*** *At any given* $\hat{\mu}_\gamma \leqslant \bar{\mu}_\gamma$*, the parent forgoes the entire calibration gain* $\frac{\delta}{2}\hat{\mu}_\gamma^2\sigma_\theta^2$*. The shortfall is continuous at* $\bar{\mu}_\gamma$ *and both regimes agree at the threshold.*

Proof. See Appendix A. □

Part (iii) captures the central mechanism of the model: *lack of bandwidth makes knowledge painful*. A parent with low bandwidth faces a larger shortfall between what accurate beliefs demand and what her constraints allow her to deliver, making the psychological cost of holding accurate beliefs higher. The same prior $\mu_{\gamma,0}$ produces more distortion in a more constrained parent, not because she is less rational, but because the cost of accurate beliefs is objectively higher for her.

3.2.2 Optimal Working Beliefs

The Stage 1 problem is:

$$\max_{\hat{\mu}_\gamma} \quad V_2(\hat{\mu}_\gamma, b, \rho) \;-\; \frac{\lambda_\gamma}{2\sigma_{\gamma,0}^2}(\hat{\mu}_\gamma - \mu_{\gamma,0})^2 \;-\; G(\hat{\mu}_\gamma, b, \rho) \tag{20}$$

The Stage 1 solution depends on the parent's reference model $(\mu_{\gamma,0}, \sigma_{\gamma,0}^2)$, which varies across parents and reflects heterogeneity in prior exposure to information about child development. We summarize the joint contribution of the reference model through the *prior ratio*:

$$\phi_0 \;\equiv\; \frac{\mu_{\gamma,0}}{\sigma_{\gamma,0}^2} \tag{21}$$

which measures the prior mean per unit of prior variance. A higher ϕ_0 reflects either a more optimistic prior mean $\mu_{\gamma,0}$ or a tighter prior $\sigma_{\gamma,0}^2$, both of which reduce the scope for downward distortion in Stage 1, the first by raising the material return to holding high beliefs, the second by raising the cost of departing from the prior mean.

To ensure the Stage 1 problem is well-posed, we impose the following assumption.

Assumption 1.

$$A_\delta \;\equiv\; \frac{\lambda_\gamma}{\sigma^2_{\gamma,0}} - \delta(\mu_\theta^2 + \sigma_\theta^2) \;>\; (1+\psi)\delta\sigma_\theta^2 \tag{22}$$

where $\psi > 0$ and $\delta > 0$. Equivalently, $\lambda_\gamma/\sigma^2_{\gamma,0} > \delta(\mu_\theta^2 + (2+\psi)\sigma_\theta^2)$.

Assumption 1 requires that the precision-weighted penalty cost $\lambda_\gamma/\sigma^2_{\gamma,0}$ exceeds δ times the expected squared cue rate $\mu_\theta^2+\sigma_\theta^2$ plus the psychological response to cue rate uncertainty $(1+\psi)\sigma_\theta^2$. It serves two roles simultaneously. First, it implies $A_\delta > 0$, which is the second-order condition for the interior regime: without it, the Stage 1 objective is convex in $\hat{\mu}_\gamma$ and unbounded above. Second, it implies the corner second-order condition holds automatically, since $\delta\mu_\theta^2 - \lambda_\gamma/\sigma^2_{\gamma,0} - \psi\delta\sigma_\theta^2 < -A_\delta - \psi\delta\sigma_\theta^2 < 0$. The condition $\sigma^2_{\gamma,0} < \lambda_\gamma/[\delta(\mu_\theta^2 + (2+\psi)\sigma_\theta^2)]$ reveals the fundamental role of prior precision: a very diffuse prior violates Assumption 1, making the interior regime structurally ill-posed regardless of other parameter values.

Proposition 4 (Optimal working belief)**.** *Under Assumption 1, let* $A_\delta \equiv \lambda_\gamma/\sigma^2_{\gamma,0} - \delta(\mu_\theta^2 + \sigma_\theta^2)$, $B \equiv \mu_\theta \ln x^R + \lambda_\gamma\phi_0$, $P \equiv A_\delta + (1+\psi)\delta\sigma_\theta^2$, *and define the* critical prior mean*:*

$$\bar{\mu}_{\gamma,0}(b,\rho) \equiv \frac{\sigma^2_{\gamma,0}}{\lambda_\gamma}\left[\bar{\mu}_\gamma(b,\rho)\,P - \mu_\theta \ln x^R\right] \tag{23}$$

The Stage 1 objective is strictly concave in $\hat{\mu}_\gamma$ *in both regimes and the unique optimal working belief is:*

$$\hat{\mu}^*_\gamma = \frac{B}{P} \qquad \text{if } \mu_{\gamma,0} \leqslant \bar{\mu}_{\gamma,0} \tag{24}$$

$$\hat{\mu}^*_\gamma = \frac{B + \sqrt{B^2 - 8A_\delta(1+\psi)\kappa(b,\rho)}}{2A_\delta} \qquad \text{if } \mu_{\gamma,0} > \bar{\mu}_{\gamma,0} \tag{25}$$

The working belief $\hat{\mu}^*_\gamma$ *is a continuous function of* $\mu_{\gamma,0}$ *everywhere: both solutions deliver* $\hat{\mu}^*_\gamma = \bar{\mu}_\gamma(b,\rho)$ *at* $\mu_{\gamma,0} = \bar{\mu}_{\gamma,0}$*. The discriminant* $B^2 - 8A_\delta(1+\psi)\kappa(b,\rho) \geqslant 0$ *if and only if* $\mu_{\gamma,0} \geqslant \bar{\mu}_{\gamma,0}$*, so the interior solution is well-defined precisely when the parent is in the interior regime.*

Proof. See Appendix A. □

Assumption 2.

$$\psi\sigma_\theta^2 \;>\; \mu_\theta^2 \tag{26}$$

Assumption 2 requires that the parent's psychological sensitivity ψ to the aspiration-capacity gap exceeds the squared signal-to-noise ratio $\mu_\theta^2/\sigma_\theta^2$ of the child's cue process. Assumption 2 is independent of Assumption 1: neither implies the other, and both may hold simultaneously or fail simultaneously depending on parameter values.

Assumption 3.

$$\bar{\mu}_\gamma \; > \; \frac{\mu_\theta \ln x^R}{P}\left[1 + \frac{\lambda_\gamma}{\sigma^2_{\gamma,0}\delta(\psi\sigma^2_\theta - \mu^2_\theta)}\right] \tag{27}$$

Assumption 3 ensures that the tension between the material gain from high beliefs and the psychological cost of deviating from prior beliefs is genuine — neither force dominates for all parents simultaneously.

Proposition 5 (Direction of distortion)**.** *Under Assumptions 1, 2, and 3, define the* corner distortion threshold*:*

$$\tilde{\mu}^{corner}_{\gamma,0} \equiv \frac{\mu_\theta \ln x^R}{\delta(\psi\sigma^2_\theta - \mu^2_\theta)} \tag{28}$$

and the interior distortion threshold*:*

$$\tilde{\mu}^{interior}_{\gamma,0} \equiv \frac{-\mu_\theta \ln x^R + \sqrt{(\mu_\theta \ln x^R)^2 + 8(1+\psi)\kappa(b,\rho)\delta(\mu^2_\theta + \sigma^2_\theta)}}{2\delta(\mu^2_\theta + \sigma^2_\theta)} \tag{29}$$

Under Assumptions 1, 2, and 3, the ordering $\tilde{\mu}^{corner}_{\gamma,0} < \bar{\mu}_{\gamma,0} < \tilde{\mu}^{interior}_{\gamma,0}$ holds, and the direction of distortion is characterized by four regions:

(i) ***Informational trap.*** *If $\mu_{\gamma,0} \leqslant \tilde{\mu}^{corner}_{\gamma,0}$, the parent is in the corner regime with upward distortion: $\hat{\mu}^*_\gamma \geqslant \mu_{\gamma,0}$. Despite distorting upward, the parent cannot escape the corner regime because her prior mean is too pessimistic for even the maximum upward distortion to push $\hat{\mu}^*_\gamma$ above $\bar{\mu}_\gamma$.*

(ii) ***Psychological trap.*** *If $\tilde{\mu}^{corner}_{\gamma,0} < \mu_{\gamma,0} \leqslant \bar{\mu}_{\gamma,0}$, the parent is in the corner regime with downward distortion: $\hat{\mu}^*_\gamma < \mu_{\gamma,0}$. Motivated reasoning keeps the parent in the corner regime: she could escape if she held beliefs equal to her prior mean, but the psychological cost of accurate beliefs drives distortion downward.*

(iii) ***Interior regime, downward distortion.*** *If $\bar{\mu}_{\gamma,0} < \mu_{\gamma,0} < \tilde{\mu}^{interior}_{\gamma,0}$, the parent is in the interior regime with downward distortion: $\bar{\mu}_\gamma < \hat{\mu}^*_\gamma < \mu_{\gamma,0}$. The parent reads child cues but holds working beliefs below her prior mean.*

(iv) ***Interior regime, upward distortion.*** *If $\mu_{\gamma,0} \geqslant \tilde{\mu}^{interior}_{\gamma,0}$, the parent is in the interior regime with upward distortion: $\hat{\mu}^*_\gamma \geqslant \mu_{\gamma,0}$. The material gain from holding high beliefs dominates the psychological cost, and the parent reads child cues while holding working beliefs above her prior mean.*

In the corner regime, the direction and magnitude of distortion are independent of bandwidth b and cue-detection expertise ρ. In the interior regime, $\tilde{\mu}^{interior}_{\gamma,0}$ is increasing in $\kappa(b,\rho)$: higher attention costs expand the downward distortion region, while higher bandwidth or cue-detection expertise shift $\tilde{\mu}^{interior}_{\gamma,0}$ leftward, expanding the upward distortion region.

Proof. See Appendix A. □

Figure 1 illustrates the four regions of Proposition 5 by plotting the optimal working belief $\hat{\mu}^*_\gamma$ against the prior mean $\mu_{\gamma,0}$. The economic implications of this structure are developed in Section 3.3.

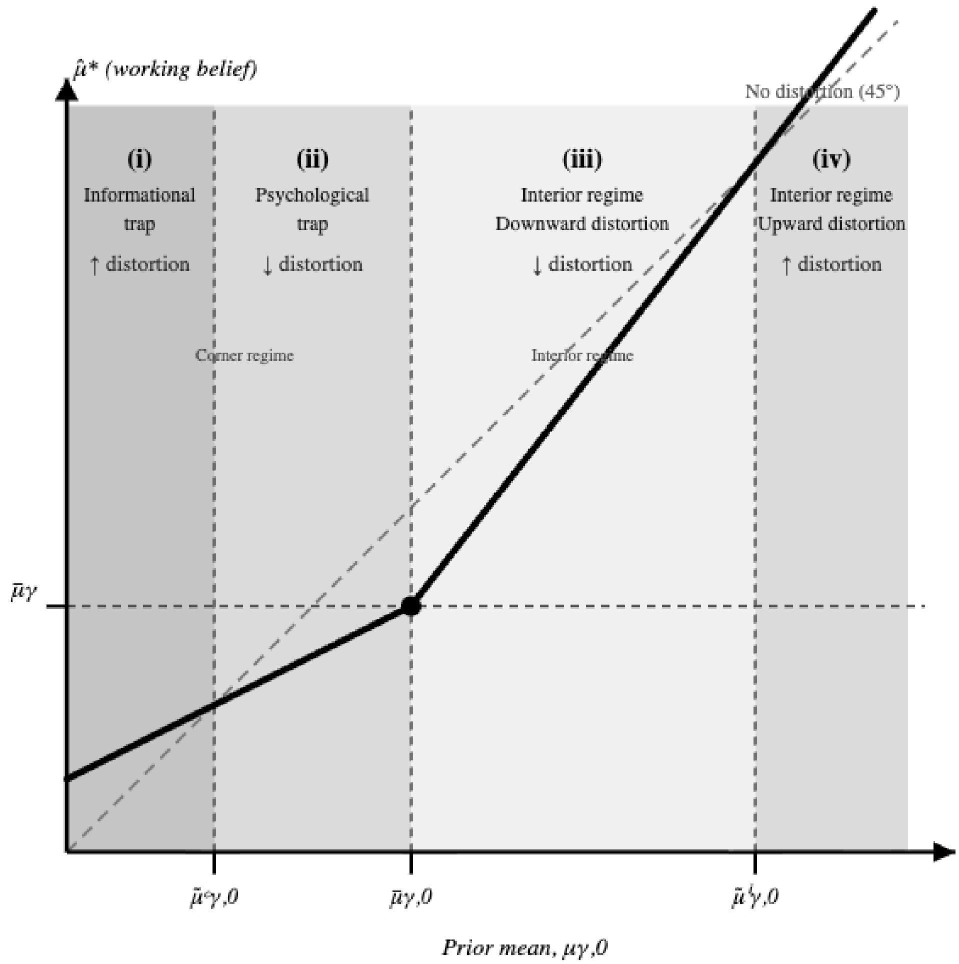

Figure 1: Belief Distortion Regions

3.3 Self-Sealing Trap

Propositions 4 and 5 identify two groups of parents. Parents with $\mu_{\gamma,0} > \bar{\mu}_{\gamma,0}$ hold working beliefs above $\bar{\mu}_\gamma$, read child cues in Stage 2, and generate calibration evidence about the returns to responsive engagement. Parents with $\mu_{\gamma,0} \leqslant \bar{\mu}_{\gamma,0}$ hold working beliefs at or below $\bar{\mu}_\gamma$, respond uniformly regardless of the child's signal, and generate no calibration evidence. The critical prior mean $\bar{\mu}_{\gamma,0}$ therefore partitions the population into those who engage responsively and those who do not. We now show that the corner regime is not merely self-fulfilling but self-sealing, a stronger and more consequential property.

To understand why, we introduce the concept of recursive complementarity across the two barriers.

Definition 1 (Recursive complementarity)**.** *The two barriers exhibit* recursive complementarity *if the marginal return to addressing one barrier is increasing in the degree to which the other barrier has been addressed. Formally, the cross-partials of the solution with respect to the policy levers governing each barrier are positive: addressing believing makes acting more productive, and addressing acting makes believing more productive.*

The complementarities established in Proposition B2 are formal instances of this property: believed returns and bandwidth are complements ($\partial^2 V_2/\partial\hat{\mu}_\gamma\,\partial b > 0$), and believed returns and cue-detection expertise are complements ($\partial^2 V_2/\partial\hat{\mu}_\gamma\,\partial\rho > 0$). In the interior regime, the two barriers reinforce each other positively: a parent who believes responsiveness matters gains more from additional bandwidth and cue-detection expertise, and a parent with greater expertise and bandwidth gains more from accurate beliefs. Recursive complementarity therefore sustains engagement in the interior regime.

The same property operates in reverse in the corner regime. A self-fulfilling trap is one in which the parent's beliefs are consistent with the behavior they induce. She believes returns are low, behaves as if they are, and the behavior generates outcomes consistent with low returns. The trap characterized in Propositions 4 and 5 is *self-sealing*: the parent's behavior not only generates evidence consistent with her distorted beliefs, but actively eliminates the evidence that would challenge them.

The distinction is important. In a merely self-fulfilling trap, the parent could in principle escape by experimenting, for example by increasing engagement and observing whether outcomes improve. The self-sealing property forecloses this. In the corner regime of Stage 2, the parent does not vary her response with child cues, so there is no variation in responses to generate identifying evidence about the returns to engagement. The parent cannot learn from experience because her behavior produces no experiential variation from which to learn. When believing fails, the working belief falls below $\bar{\mu}_{\gamma,0}$ and the parent is placed in the Stage 2 corner. When acting fails, the parent responds uniformly, no calibration evidence is generated, and the experiential feedback that would raise the penalty cost of distortion in Stage 1 is eliminated. Each failure reinforces the other, and the system is stable at the corner.

3.4 Escaping the Trap

Propositions 5 and B3 identify the parameters that govern whether a parent is in the trap and how far she is from escaping it. The trap is broken when $\hat{\mu}^*_\gamma$ crosses $\bar{\mu}_\gamma$ from below — either because the working belief rises, or because the threshold falls, or both. The model identifies four distinct channels through which this can occur, each targetable by a different type of intervention.

The **belief channel** raises $\mu_{\gamma,0}$, moving the parent's prior mean toward and above the critical prior mean $\bar{\mu}_{\gamma,0}$ (Proposition B3(i)). Since $\bar{\mu}_{\gamma,0}$ is independent of $\mu_{\gamma,0}$, this channel operates entirely on the parent's side — it shifts the parent toward the boundary rather than shifting the boundary toward the parent. In the corner regime, a unit increase in $\mu_{\gamma,0}$ raises the working belief by $(\lambda_\gamma/\sigma^2_{\gamma,0})/P < 1$ — less than one-for-one, because the accuracy cost disciplines but does not eliminate distortion. This channel is most effective for Region 2 parents, whose prior mean is close to $\bar{\mu}_{\gamma,0}$ and requires only a moderate shift to escape the psychological trap. For Region 1 parents, the belief channel must raise $\mu_{\gamma,0}$ above both $\tilde{\mu}^{\text{corner}}_{\gamma,0}$ and $\bar{\mu}_{\gamma,0}$ (Proposition 5(i)–(ii)).

The **bandwidth channel** raises b, lowering $\bar{\mu}_{\gamma,0}$ at rate $\sigma^2_{\gamma,0}P\bar{\mu}_\gamma/(2b\lambda_\gamma)$ and simultaneously reducing $\kappa(b,\rho)$ (Proposition B3(ii)). The reduction in κ has two effects: it lowers the cue detection threshold $\bar{\mu}_\gamma$, making the interior regime more accessible, and it reduces the psychological cost $G(\hat{\mu}_\gamma, b, \rho)$ of holding accurate beliefs in the interior regime. Higher bandwidth therefore makes responsive engagement less costly and accurate beliefs less painful simultaneously.

The **cue-detection expertise channel** raises ρ, lowering $\bar{\mu}_{\gamma,0}$ at rate $\sigma^2_{\gamma,0}P\bar{\mu}_\gamma/[2(1+\rho)\lambda_\gamma]$ and simultaneously reducing $\kappa(b,\rho)$, with the same dual effect as the bandwidth channel (Proposition B3(iii)). The relative effectiveness of the bandwidth and cue-detection expertise channels is $(1+\rho)/b$: bandwidth is relatively more effective when ρ is high and b is low, and cue-detection expertise is relatively more effective when b is high and ρ is low. When the two channels are equally effective, the optimal intervention targets whichever is less costly to activate.

The **trap compression channel** raises $\ln x^R$, lowering $\bar{\mu}_{\gamma,0}$ at rate $\sigma^2_{\gamma,0}\mu_\theta/\lambda_\gamma$ and simultaneously raising the working belief of corner regime parents at rate μ_θ/P (Proposition B3(iv)). Unlike the bandwidth and cue-detection expertise channels, the trap compression channel does not operate through $\kappa(b,\rho)$ and does not affect the psychological cost function G in either regime. It changes which regime the parent occupies and raises her working belief without changing either the psychological cost of accurate beliefs or her prior mean. This is the channel through which improvements in the baseline environment — better materials, structured activities, enriched physical surroundings — compress the trap and improve child outcomes without necessarily shifting parental beliefs about the returns to responsive interaction.

These channels are neither mutually exclusive nor jointly necessary. Any single channel, if sufficiently activated, can move the parent from the corner to the interior regime. Different programs activate different channels at different intensities, and the model provides a unified language for characterizing what each program is doing and under what conditions it is sufficient. Once the parent enters the interior regime through any route, she begins reading child cues and the child's language environment improves.

4 Intervention

The LENA Start Program is a ten-week group-based parenting intervention designed to increase the quantity and quality of parent-child conversational interactions in the first three years of life.

The program operates through five components (Education, Coaching, Feedback, Book Reading, and Reporting), each of which targets specific parameters in the model developed in Section 3. We present a more detailed explanation of the intervention components in Appendix C.1. In addition, a session-by-session description of the program is provided in Appendix C.2.

This section maps each component onto the model's channels and derives testable implications the model generates for the program's effects.

4.1 The Five Components and Their Model Counterparts

Education. The Education component delivers scientific content about the returns to responsive interaction through videos and a parent handbook (Appendix C.1). It operates through the **belief channel**: by providing information about γ passively and in a group setting, Education raises $\mu_{\gamma,0}$, collapses $\sigma^2_{\gamma,0}$, and increases A_δ, making downward distortion of $\hat{\mu}^*_\gamma$ more costly.

Coaching. The Coaching component teaches parents to recognize and respond to child communicative bids and builds a repertoire of interaction activities integrated into daily routines (Appendix C.1). It operates through two channels simultaneously. Through the **cue-detection expertise channel**, Coaching raises ρ, reducing $\kappa(b,\rho)$, lowering $\bar{\mu}_\gamma$ and $\bar{\mu}_{\gamma,0}$, and reducing the psychological cost G. Through the **trap compression channel**, it raises x^R via repertoire building, shifting the optimal response upward in both regimes and compressing the trap region $(\tilde{\mu}^{\text{corner}}_{\gamma,0}, \bar{\mu}_{\gamma,0}]$.

Feedback. The Feedback component provides weekly personalized LENA reports reviewed with a Family Liaison (Appendix C.4). It operates through the **trap compression channel** by anchoring x^R to quantitative targets for conversational turns, and through the **cue-detection expertise channel** by building attentiveness to conversational behavior across recording sessions.

Book Reading. The Book Reading component provides age-appropriate books and teaches dialogic reading techniques (Appendix C.1). It operates through the **cue-detection expertise channel** by giving parents a structured context in which to practice reading child communicative bids, raising ρ in the reading context specifically and lowering $\bar{\mu}_\gamma$.

Reporting. The Reporting component administers a monthly Developmental Snapshot providing the parent with her child's language development percentile (Appendix C.1). It operates through the **belief channel** by providing output-side evidence about the consequences of engagement — a signal about γ that is harder to rationalize away than general scientific content, shifting $\mu_{\gamma,0}$ upward.

Table 1 summarizes the mapping of each component onto the model's parameters and channels.

Table 1: LENA Start: Components, Parameters, and Channels

Component	Stage	Parameter moved	Channel
Education	1	$\mu_{\gamma,0} \uparrow$, $\sigma^2_{\gamma,0} \downarrow$, $A_\delta \uparrow$, $D_\gamma \uparrow$	Belief
Coaching	1, 2	$\rho \uparrow$, $x^R \uparrow$, $G \downarrow$, $\bar{\mu}_\gamma \downarrow$, $\bar{\mu}_{\gamma,0} \downarrow$	Cue-detection expertise, trap compression
Feedback	1, 2	$\rho \uparrow$, $x^R \uparrow$	Cue-detection expertise, trap compression
Book Reading	2	$\rho \uparrow$	Cue-detection expertise
Reporting	1	$\mu_{\gamma,0} \uparrow$	Belief

Notes: This table maps the LENA Start program components to the structural parameters and theoretical channels discussed in Section [X]. Upward arrows (↑) indicate an increase in the parameter value; downward arrows (↓) indicate a decrease.

4.2 Testable Implications

The model generates four testable implications that distinguish it from alternative accounts of parenting behavior and intervention effectiveness. Appendix C.5 provides the full derivations.

TI-1: Large effects on parental beliefs. LENA Start activates both the belief channel (Education and Reporting raise $\mu_{\gamma,0}$ and collapse $\sigma^2_{\gamma,0}$, increasing the accuracy cost of distortion) and the cue-detection expertise channel (Coaching and Feedback raise ρ, reducing G and making accurate beliefs less painful). A program that activates both channels simultaneously produces larger and more persistent belief effects than one that activates only one. The model predicts large effects on the belief factor score.

TI-2: Effects on responsive engagement. The intervention increases responsive engagement through two margins: moving corner regime parents into the interior regime (via the belief and cue-detection expertise channels), and raising engagement intensity for parents already in the interior regime (since ω^*_θ is decreasing in $\hat{\mu}^*_\gamma$). Both margins predict positive treatment effects on conversational turns and child-initiated exchanges.

TI-3: Differential effects on behavioral versus material components of the home environment. The model distinguishes two components of parental investment: the cue-contingent component x, which varies with the child's signal and reflects responsive engagement, and the reference investment x^R, which captures the parent's default level of engagement regardless of the child's cue. LENA Start activates multiple channels simultaneously. Therefore, the model predicts positive treatment effects on both the cue-contingent and level behavioral components of the home environment.

A sharper prediction follows from what the program does not do: it provides no resources for parents to purchase learning materials or toys. Subscales of a home environment measure that reflect access to materials rather than parental behavior should therefore be unaffected. This creates a within-instrument falsification test embedded inside the home environment measure itself. A program that operates only through the belief channel would raise only the cue-contingent behavioral component; one that operates only through the trap compression channel would raise only the level behavioral component; one that provides material resources would raise the materials component. A multi-channel program with no material resource component predicts a specific pattern: behavioral subscales respond, the materials subscale does not.

TI-4: Null effects on preferences and self-efficacy. The psychological cost G is belief-specific, not a general preference parameter or global confidence measure. The intervention narrows the aspiration-capacity gap through cue-detection expertise, not by changing what the parent values or how confident she feels in general. The model therefore predicts null effects on global preferences and general self-efficacy, distinguishing it from accounts based on motivation or confidence.

5 Experimental Design

5.1 Setting

The study was conducted in the Alief Independent School District (AISD), a large urban district in Houston, Texas. Alief ISD serves a predominantly Hispanic population with substantial economic disadvantage.[8]

Program delivery was carried out by AISD Family Liaisons. They are permanent staff members assigned to each school who oversee activities with families and students. Systematic fidelity documentation confirmed high adherence and competence levels across language groups and cohorts, minimizing concerns about variability in program delivery (see Appendix C.6).

5.2 Randomization Protocol

Eligibility and recruitment. Families living in the Alief ISD catchment area were invited to participate through postal mail advertising by the AISD Family and Community Engagement department. Inclusion criteria required the primary family member (PFM) to be 18 years or older, have a child between 16 and 32 months old, and speak either Spanish or English as their native language. Interested families were screened for eligibility by AISD liaisons and then scheduled for a consent visit with the Rice University research team, during which the LENA Start Program, the program lottery, and the evaluation study were explained in detail.

Randomization. Participants were stratified by language group (English and Spanish) and cohort and randomly assigned approximately equally to treatment or control within each stratum. Treatment group families participated in both the LENA Start Program and the evaluation study. Control group families participated only in the evaluation study and received books to reduce attrition.

[8] See Cunha, Hu, Salvati, Wolpin, and Zeng (2025) for an in-depth discussion about the families in the Alief ISD catchment area.

Sample. Participants were recruited and enrolled over three academic years, divided into six distinct cohorts: Fall 2022, Spring 2023, Fall 2023, Spring 2024, Fall 2024, and Spring 2025. In total, 293 families were successfully enrolled — 139 English-speaking and 154 Spanish-speaking. Figure D1 illustrates the randomization process. Detailed enrollment counts across cohort-language groups are presented in Table D1, and Table C4 summarizes session attendance and graduation rates, averaging approximately 75% across all cohorts.

Compliance and attrition. Both treatment and control group participants completed a comprehensive baseline survey and submitted LENA recordings at two points: enrollment (baseline) and approximately ten weeks later (endline). Across all cohorts, 203 participants provided valid 12-hour recordings at both baseline and endline, forming the primary analytic sample for LENA-based outcomes.

5.3 Measurement

Data collection was administered by the Rice University research team at baseline and endline. It assessed outcomes corresponding to each of the model's testable implications, as well as baseline control variables characterizing the child and primary caregiver. Further details about variable construction are provided in Appendix E.

TI-1: Stage 1 outcomes — beliefs and knowledge. Parental beliefs were measured as in (Cunha, Gerdes, Hu, and Nihtianova 2024) and parental knowledge using the questionnaire developed by Suskind et al. (2016). To account for potential measurement error, we followed Cunha, Elo, and Culhane (2022) and conducted a factor analysis on the items in both questionnaires. The resulting factor scores provide an error-adjusted measure of parental beliefs about the returns to responsive interaction.

TI-2: Stage 2 outcomes — cue detection and responsive engagement. The LENA recording system documented each child's language environment over a typical day at baseline and endline (Gilkerson and Richards 2020). Families received specialized clothing with a vest pocket to securely hold the LENA Digital Language Processor, enabling automated capture of adult word counts, child vocalizations, and conversational turns. Technical details of how conversational turns are counted and how recording periods are selected are provided in Appendix E.[9]

The primary outcome measure, conversational turns, maps directly onto the Stage 2 engagement precision ω_θ^*: more conversational turns reflect a lower ω_θ^* and more attentive monitoring of child cues. The model predicts positive treatment effects on conversational turns through two margins: parents in the corner regime moving into the interior regime and activating cue detection, and parents already in the interior regime engaging more intensively as $\hat{\mu}_\gamma^*$ rises (TI-2).

We utilized the Advanced Data Extractor (ADEX) feature of the LENA system to decompose

[9] A few parents provided recordings for two days. For those parents, one recording was randomly selected for analysis.

conversational turns by initiator: Key Child, Adult Female, or Adult Male. ADEX groups time-adjacent LENA segments into conversation blocks and identifies the initiator by scanning forward from the block's start time to the first qualifying live-speech segment. Each block is characterized by its initiator, its duration, and the number of conversational turns it contains. Technical details of the ADEX algorithm are provided in Appendix E.

This initiator decomposition is what allows ADEX to go beyond raw conversational turns as a measure of responsive parenting. A raw turn count aggregates exchanges regardless of who started them, conflating two behaviorally distinct phenomena: a parent who reads a child's cue and responds to it, and a parent who initiates an exchange unprompted. Child-initiated blocks are the behavioral signature of a parent operating in the interior regime: they arise when the parent is reading child cues and calibrating her response to them, generating $\mathrm{Cov}(x, \theta) > 0$. Adult-initiated blocks reflect level investment independent of the child's signal. In the corner regime, $\mathrm{Cov}(x, \theta) = 0$, so child-initiated blocks are absent even if adult-initiated blocks are not.

TI-3: Behavioral versus material components of the home environment. The StimQ Toddler questionnaire (Cates et al. 2023) assesses the cognitive home environment across four subscales, each mapping onto a distinct element of the model. The StimQ provides an independent, survey-based measure of parental responsivity that complements the LENA system: whereas LENA captures the acoustic record of adult-child interaction, the StimQ captures the caregiver's self-reported engagement behaviors. The two instruments triangulate responsive parenting from different angles and agreement between them strengthens the interpretation of treatment effects on TI-2 and TI-3.

Parental Verbal Responsivity (PVR) maps onto the cue-contingent component of investment x; Reading (READ) and Parental Involvement in Developmental Advance (PIDA) map onto the reference investment x^R through the level channel; and Availability of Learning Materials (ALM) reflects the material component of the home environment. Since LENA Start provides no resources for parents to purchase learning materials, ALM serves as a within-instrument placebo. The model predicts positive treatment effects on PVR, READ, and PIDA, and a null effect on ALM. Full subscale descriptions are provided in Appendix E.

TI-4: Placebo outcomes — global preferences, self-efficacy, and social support. The model predicts null treatment effects on outcomes reflecting general preferences or psychological traits rather than parenting-specific beliefs and constraints. Parental self-efficacy and social support are measured using scales developed by the LENA Foundation (Cunha, Gerdes, Hu, and Nihtianova 2024). Global preferences — altruism, reciprocity, trust, and patience — are measured using the validated instrument of Falk et al. (2018). Non-zero effects on these placebo outcomes would be inconsistent with the model and would suggest alternative mechanisms based on confidence, motivation, or social network effects. The placebo tests therefore provide a direct falsification exercise for the model's mechanism.

5.4 Summary Statistics

Tables D2, D3, and D4 report summary statistics and balance tests organized by testable implication. Table D2 covers demographic control variables and TI-1 belief outcomes. Table D3 covers TI-2 responsive engagement outcomes from LENA and ADEX. Table D4 covers TI-3 StimQ components and TI-4 placebo outcomes.

Panel A of Table D2 confirms demographic balance between treatment and control groups across all key covariates, validating the randomization. The sample is predominantly Hispanic (66–70%), with approximately half of primary caregivers having attended college and more than 70% living with a co-parent. Approximately 91% of primary family members are mothers. The mean child age is approximately 23 months, with roughly equal gender distributions across groups. All p-values in Panel A exceed 0.10. Panel B of Table D2 reports baseline balance on TI-1 outcomes (parental knowledge, parental beliefs, and the belief factor score) with p-values of 0.139, 0.284, and 0.583 respectively, indicating no pre-existing differences in parental beliefs or knowledge.

Table D3 reports baseline balance on TI-2 outcomes for the 203 families with valid baseline LENA recordings. Conversational turns, adult word counts, and all ADEX-decomposed measures of child-initiated, female adult-initiated, and male adult-initiated conversation blocks are balanced across treatment and control groups, with all p-values exceeding 0.10.

Panel A of Table D4 reports balance on TI-3 StimQ outcomes for the full sample of 238 families. All subscale scores (StimQ total, PVR, READ, PIDA, and ALM) are balanced across groups, with p-values ranging from 0.301 to 0.862. Panel B reports balance on TI-4 placebo outcomes, with sample sizes ranging from 177 to 238 depending on the outcome. Two variables approach but do not reach conventional significance levels: Parenting Social Support ($p = 0.066$) and Trust ($p = 0.088$). All other placebo outcomes are well-balanced. Neither variable survives a Bonferroni correction for multiple testing across the eight placebo outcomes, which requires $p < 0.0125$ for significance at the 10% level, and we return to both when discussing the TI-4 attrition robustness results in Appendix G.

5.5 Empirical Model

To estimate the intent-to-treat (ITT) effects of the LENA Start Program, we employ regression models specified as follows.

For analyses involving LENA recording outcomes:

$$Y_{i,1} = \beta_0 + \beta_1 Z_i + \beta_2 Y_{i,0} + X_i \beta_3 + \sum_{t=0}^{1} \beta_{4+t} R_{i,t} + \sum_{j=2}^{J} \tau_j B_{i,j} + \epsilon_{i,1} \tag{30}$$

For survey-based outcomes:

$$Y_{i,1} = \beta_0 + \beta_1 Z_i + \beta_2 Y_{i,0} + X_i \beta_3 + \sum_{j=2}^{J} \tau_j B_{i,j} + \epsilon_{i,1} \tag{31}$$

Here $Y_{i,1}$ is the outcome at endline, $Y_{i,0}$ the corresponding baseline measure, and Z_i the binary treatment indicator ($Z_i = 1$ for treatment, $Z_i = 0$ for control). Covariates X_i encompass demographic characteristics of the child and primary caregiver, including language group, child gender and age, caregiver education, race, co-parenting status, and marital status. Cohort indicators $B_{i,j}$ control for systematic variation across the six cohorts, with the first cohort as the baseline reference. For LENA outcomes, $R_{i,t}$ incorporates recording-specific controls including start times and weekend recording indicators.

All dependent variables are standardized to mean zero and standard deviation one, facilitating direct comparison across outcomes. The estimated coefficient β_1 captures the standardized ITT effect of treatment assignment. We report robust standard errors, standard errors clustered at the cohort-language level, and randomization inference p-values (Young 2019).

6 Results

We report the intent-to-treat (ITT) effects of the LENA Start Program on five categories of outcomes, each corresponding to a testable implication derived in Section 4.2. The results are organized to mirror the model's structure: we begin with the belief outcomes (TI-1), proceed to the engagement outcomes including conversational turns and the ADEX decomposition (TI-2), document the differential effects on behavioral versus material components of the home environment (TI-3), and report the null effects on placebo outcomes (TI-4). The final subsection summarizes the sensitivity and attrition analyses, as well as effect heterogeneity.

6.1 TI-1: Effects on Beliefs

Table 2 reports the ITT effects on parental knowledge and beliefs. Random assignment to treatment increases parental knowledge by nearly 50% of a standard deviation ($p < 0.01$), parental beliefs by 69% of a standard deviation ($p < 0.01$), and the belief factor score by 86% of a standard deviation ($p < 0.01$).

These results are consistent with TI-1. The model predicts belief effects for a program that activates both the belief channel (Education raising $\mu_{\gamma,0}$ and collapsing $\sigma^2_{\gamma,0}$, increasing the accuracy cost of downward distortion) and the cue-detection expertise channel (Coaching and Feedback raising ρ and reducing G, making accurate beliefs less painful to hold).

Table 2: Impact of the LENA Start Program on Parental Measures

	(1) Baseline	(2) +Child	(3) Full Controls
Parental Knowledge	0.486***	0.487***	0.496***
	(0.095)	(0.097)	(0.097)
	[0.081]	[0.082]	[0.086]
	{0.000}	{0.000}	{0.000}
Parental Beliefs	0.706***	0.696***	0.694***
	(0.118)	(0.118)	(0.118)
	[0.131]	[0.128]	[0.126]
	{0.000}	{0.000}	{0.000}
Belief Factor Score	0.833***	0.839***	0.866***
	(0.110)	(0.113)	(0.112)
	[0.100]	[0.103]	[0.098]
	{0.000}	{0.000}	{0.000}
Child Controls	No	Yes	Yes
PFM Controls	No	No	Yes

Notes: $N = 238$. Robust standard errors in $(\cdot)$, clustered standard errors in $[\cdot]$, and Randomization Inference (RI) p-values in $\{\cdot\}$. All models include baseline outcome, cohort, and language group dummies. Child controls: male, age. PFM controls: education, mother, co-parenting, ethnicity, and marital status. * p<0.10, ** p<0.05, *** p<0.01.

6.2 TI-2: Effects on Responsive Engagement

6.2.1 Conversational Turns

Table 3 reports the ITT effects on LENA recording outcomes. The preferred specification, which controls for recording start time, weekend indicators, and demographics of the child and primary family member, yields an ITT effect on conversational turns of 30% of a standard deviation ($p < 0.01$). The effect is robust across model specifications, ranging from nearly 30% to slightly above 32% of a standard deviation. Adult word count is weakly significant at baseline but loses significance with the addition of demographic controls, consistent with TI-2: the intervention raises the quality of interaction, cue-contingent responses, rather than simply the quantity of adult speech.

6.2.2 ADEX Decomposition

Table 4 reports the ITT effects on conversational blocks decomposed by initiator using the ADEX feature of the LENA system. Child-initiated conversational blocks increase by about 26% of a standard deviation ($p < 0.05$), with associated conversational turns increasing by over 30% of a standard deviation, block duration increasing by nearly 35%, and adult word count within child-initiated blocks increasing by about 34% of a standard deviation. By contrast, blocks initiated by adult males show no significant effects. This is expected: no fathers were enrolled in the program, so adult male speech in the LENA recordings reflects household members who were not exposed to the intervention. Blocks initiated by adult females show effects of comparable magnitude to

Table 3: Impact of LENA Start Program on Adult Words and Conversational Turns

	Adult Words			Conversational Turns		
	(1)	(2)	(3)	(1)	(2)	(3)
Treatment	0.214*	0.173	0.180	0.326***	0.298***	0.300***
	(0.120)	(0.121)	(0.123)	(0.101)	(0.095)	(0.099)
	[0.081]	[0.090]	[0.102]	[0.061]	[0.061]	[0.068]
	{0.091}	{0.165}	{0.157}	{0.002}	{0.003}	{0.003}
Start Time / Weekend	No	Yes	Yes	No	Yes	Yes
Child Controls	No	Yes	Yes	No	Yes	Yes
PFM Controls	No	No	Yes	No	No	Yes

Notes: $N = 203$. Robust standard errors in $(\cdot)$, clustered standard errors in $[\cdot]$, and RI p-values in $\{\cdot\}$. All specifications include baseline dependent variables, language group, and cohort dummies. Child controls include gender and age. PFM controls include education, mother dummy, co-parenting, ethnicity, and marital status. * p<0.10, ** p<0.05, *** p<0.01.

child-initiated blocks for turns, duration, and word count, consistent with the concentration of treatment effects among the primary family members who participated in LENA Start.

Table 4: Impact of LENA Start Program on Conversation Block Metrics

	(1) Initiated by Child	(2) Initiated by Female	(3) Initiated by Male
Conv. Block Count	0.262**	0.105	0.160
	(0.123)	(0.121)	(0.134)
	[0.102]	[0.114]	[0.114]
	{0.033}	{0.398}	{0.220}
Conv. Turn Count	0.308***	0.267**	0.170
	(0.107)	(0.108)	(0.129)
	[0.065]	[0.077]	[0.112]
	{0.007}	{0.021}	{0.175}
Block Duration	0.347***	0.271**	0.135
	(0.107)	(0.120)	(0.138)
	[0.079]	[0.097]	[0.113]
	{0.002}	{0.030}	{0.308}
Adult Word Count	0.339**	0.274**	0.048
	(0.110)	(0.126)	(0.143)
	[0.081]	[0.099]	[0.096]
	{0.003}	{0.035}	{0.711}

Notes: Robust standard errors in $(\cdot)$, clustered standard errors in $[\cdot]$, and Randomization Inference (RI) p-values in $\{\cdot\}$. All specifications include controls for baseline dependent variables, cohort, language groups, and demographic characteristics. * p<0.10, ** p<0.05, *** p<0.01.

The ADEX decomposition provides a direct measure of the cue-contingent component of parental investment introduced in Section 5. Child-initiated blocks are the behavioral signature of a parent operating in the interior regime: they arise when the parent is reading child cues and calibrating her response to them, generating $\mathrm{Cov}(x, \theta) > 0$. The concentration of treatment effects in child-initiated blocks — rather than adult-initiated blocks — is therefore evidence that the intervention shifted the nature of parental investment from uniform to responsive, not merely its level. It also mitigates concerns about parental performance effects: if the observed gains in conversational turns were driven primarily by parents performing better on recording days, we would expect adult-initiated blocks to show larger effects than child-initiated blocks. The data show the opposite pattern.

6.3 TI-3: Effects on Behavioral vs. Material Components of Early Investments

Table 5 reports the ITT effects on the StimQ Toddler questionnaire, both overall and by subscale. Treatment assignment increases total StimQ scores by 58.3% of a standard deviation ($p < 0.01$), reflecting a broad improvement in the quality of the home language environment and parental engagement. The results confirm TI-3. All three behavioral subscales show significant positive effects: READ increases by 70.8% of a standard deviation ($p < 0.01$), PVR by 45.6% ($p < 0.01$), and PIDA by 28.2% ($p < 0.05$). The materials subscale, ALM, is statistically indistinguishable from zero (18.3%, $p > 0.10$), consistent with the null prediction for a subscale that reflects access to purchased learning materials rather than parental behavior.

Table 5: Impact of LENA Start Program on the Home Environment (StimQ)

	(1) Baseline	(2) +Child	(3) Full Controls
StimQ Total Score	0.574***	0.568***	0.583***
	(0.091)	(0.092)	(0.093)
	[0.089]	[0.092]	[0.090]
	{0.000}	{0.000}	{0.000}
Reading (READ)	0.699***	0.693***	0.708***
	(0.099)	(0.099)	(0.100)
	[0.120]	[0.123]	[0.122]
	{0.000}	{0.000}	{0.000}
PIDA	0.250**	0.252**	0.282**
	(0.105)	(0.105)	(0.104)
	[0.109]	[0.110]	[0.102]
	{0.014}	{0.003}	{0.007}
PVR	0.453***	0.440***	0.456***
	(0.103)	(0.104)	(0.105)
	[0.092]	[0.091]	[0.089]
	{0.000}	{0.000}	{0.000}
ALM	0.177	0.186	0.183
	(0.109)	(0.110)	(0.111)
	[0.106]	[0.108]	[0.104]
	{0.100}	{0.085}	{0.102}
Child Controls	No	Yes	Yes
PFM Controls	No	No	Yes

Notes: $N = 238$. Robust standard errors in $(\cdot)$, clustered standard errors in $[\cdot]$, and Randomization Inference (RI) p-values in $\{\cdot\}$. PIDA: Parental Involvement in Developmental Advance; PVR: Parental Verbal Responsivity; ALM: Availability of Learning Materials. * p<0.10, ** p<0.05, *** p<0.01.

The ordering of effect sizes across the three behavioral subscales maps onto the model's channel structure. The READ effect is the largest, consistent with LENA Start dedicating an entire program component (Book Reading) to building shared reading routines, activating both the trap compression channel through repertoire building and the cue-detection expertise channel through dialogic reading techniques. PVR, which captures contingent verbal engagement calibrated to the child's communicative bids, rises through the belief and cue-detection expertise channels simultaneously, producing a larger effect than PIDA. PIDA, which captures intentional teaching activities such as counting and naming colors, is activated primarily through the trap compression channel and shows the smallest effect among the three behavioral subscales. The within-instrument contrast between the three behavioral subscales and ALM constitutes a tight falsification test: it rules out general program enthusiasm or social desirability as the driver of the behavioral effects, since those mechanisms would have raised ALM alongside the other subscales.

6.4 TI-4: Null Effects on Placebo Outcomes

Table 6 reports the ITT effects on placebo outcomes across two categories: self-efficacy and social support, and global preferences. The results are consistent with TI-4.

The contrast with the preceding results is sharp. The intervention produces large and precisely estimated effects on the theoretically relevant outcomes: nearly 87% of a standard deviation on parental beliefs (TI-1), 30% on conversational turns (TI-2), and 71%, 46%, and 28% on the three behavioral StimQ subscales (TI-3). Against this backdrop, the placebo outcomes are uniformly small and imprecise. Self-efficacy shows a borderline effect of 17.5% of a standard deviation under robust standard errors ($p < 0.10$). Social support is indistinguishable from zero under all specifications. Among global preferences, measured using validated instruments from Falk et al. (2018), only negative reciprocity reaches conventional significance under clustered standard errors, and even this result is fragile given the high rate of missing data for global preference questions.

This pattern is an important distinguishing feature of the model. The intervention is not a general empowerment program: if it were, we would expect positive effects on self-efficacy and global preferences of comparable magnitude to the belief and engagement effects. Instead, the data show large effects precisely where the model predicts them (beliefs, responsive engagement, and behavioral components of the home environment) and much smaller effects everywhere the model predicts null. The model predicts that the intervention reduces the psychological cost G of the specific shortfall between parenting aspiration and constrained capacity, without affecting general utility function parameters. The data support the model's mechanism over alternatives based on confidence, motivation, or general empowerment.

6.5 Robustness and Heterogeneity

In this section, we briefly discuss the robustness of the main results along two dimensions: sensitivity to recording duration and attrition. We also summarize our findings about treatment effect heterogeneity.

Table 6: Impact of LENA Start Program on Placebo Outcomes

	(1) Baseline	(2) +Child	(3) Full Controls
Parenting Efficacy	0.180*	0.181*	0.175*
	(0.096)	(0.097)	(0.096)
	[0.083]	[0.084]	[0.091]
	{0.057}	{0.057}	{0.065}
Parenting Support	0.160	0.169	0.168
	(0.102)	(0.103)	(0.101)
	[0.095]	[0.096]	[0.100]
	{0.129}	{0.115}	{0.113}
Altruism	0.187	0.187	0.186
	(0.121)	(0.121)	(0.125)
	[0.136]	[0.132]	[0.139]
	{0.149}	{0.151}	{0.168}
Pos. Reciprocity	0.070	0.083	0.093
	(0.131)	(0.129)	(0.134)
	[0.162]	[0.156]	[0.169]
	{0.577}	{0.508}	{0.474}
Neg. Reciprocity	-0.243*	-0.245*	-0.276**
	(0.125)	(0.125)	(0.126)
	[0.087]	[0.088]	[0.084]
	{0.050}	{0.049}	{0.036}
Trust	0.152	0.153	0.143
	(0.115)	(0.116)	(0.117)
	[0.118]	[0.117]	[0.116]
	{0.191}	{0.188}	{0.227}
Patience	0.029	0.039	0.004
	(0.129)	(0.130)	(0.129)
	[0.109]	[0.108]	[0.125]
	{0.830}	{0.770}	{0.978}

Notes: Robust standard errors in $(\cdot)$, clustered standard errors in $[\cdot]$, and Randomization Inference (RI) p-values in $\{\cdot\}$. Sample size N ranges from 177 to 238. Full controls include lagged dependent variables, cohort/language dummies, and demographic characteristics. * p<0.10, ** p<0.05, *** p<0.01.

Sensitivity to recording duration (Appendix F). The main results use the best 12-hour LENA recording period. Appendix F replicates all analyses using the best 10-hour and 8-hour periods. The conversational turns effect is robust across all durations. Adult word count loses significance at shorter durations, consistent with its marginal significance in the main specification. The concentration of effects in child-initiated ADEX blocks persists across all recording windows, confirming that the shift from uniform to responsive parenting is not an artifact of the recording duration chosen.

Attrition (Appendix G). Of the 293 enrolled families, 238 provided survey data at both baseline and endline and 203 provided valid 12-hour recordings. For both survey attrition ($p = 0.07$) and recording attrition ($p = 0.04$), the control group shows higher attrition, with selective attrition on observable characteristics concentrated in the control group on English language and baseline StimQ score.

All significant treatment effects are robust to inverse probability weighting and attrition balancing. The belief factor score, parental knowledge, READ, PVR, and PIDA retain significance across all five specifications including the Lee lower bound, providing the strongest available evidence against an attrition-driven explanation of the main results. Parental beliefs retain significance under IPW and balancing but have insignificant Lee lower bounds, consistent with the mechanical consequence of trimming the treatment group under worst-case selection on unobservables. Conversational turns and the child-initiated ADEX outcomes are robust under IPW and balancing but have insignificant Lee lower bounds, which is consistent with the pattern of attrition in the data: recording attrition is higher in the control group, so the Lee bounds trim the treatment group. The within-instrument falsification result for TI-3 becomes weakly significant under IPW and significant at $p < .05$ under balancing, but this result is sensitive to the specification of the probit model we use for IPW and balancing. The TI-4 placebo outcomes show no consistent significant effects across corrections, with the isolated significant results for social support being weakly significant across specifications. Negative reciprocity is consistently significant at $p < 0.05$, although the severe missingness in the data ($N = 177$) makes interpretation uncertain.

Heterogeneity (Appendix H). We estimate Conditional Average Treatment Effects using Generalized Random Forests (Athey, Tibshirani, and Wager 2019). Across all demographic subgroups (including education, income, language, child age, baseline parental knowledge, baseline beliefs, and baseline StimQ) confidence intervals overlap substantially, providing no evidence of heterogeneous treatment effects. Two explanations are consistent with this finding. First, the model predicts that heterogeneity in treatment effects is driven by b and ρ, which may not vary significantly by observable demographic characteristics in our sample, given that Alief ISD serves an economically disadvantaged population with possibly limited variation in bandwidth constraints across families. Second, the sample size of 293 enrolled families may not provide sufficient statistical power to detect heterogeneous treatment effects even if they exist. We cannot distinguish between these two explanations with the current data.

7 Discussion

What Do Belief Elicitation Instruments Measure? The belief elicitation procedure used in this paper and in the broader literature asks parents hypothetical questions about the returns to responsive interaction. The model raises a subtle identification question: does this procedure elicit the prior belief $\mu_{\gamma,0}$ or the working belief $\hat{\mu}^*_{\gamma}$? The two are distinct objects in the model. The prior is the parent's reference model before Stage 1 distortion, while the working belief is what governs her behavior after motivated reasoning has operated. Because the elicitation question is hypothetical and removed from the behavioral context that triggers motivated reasoning, it may capture something closer to $\mu_{\gamma,0}$ than to $\hat{\mu}^*_\gamma$. A related problem arises from social desirability bias: Case 2 and Case 3 parents, whose working beliefs have been distorted downward by motivated reasoning, may report something closer to their prior when surveyed because they might expect it to be the socially desirable answer. Both forces imply that the survey measure systematically overstates $\hat{\mu}^*_\gamma$ for downward distorted parents at baseline, attenuating the observed correlation between reported beliefs and investments relative to the true relationship between working beliefs and investments. If this conjecture is true, two implications follow. First, the treatment effect on the belief factor score (87% of a standard deviation) is a lower bound on the intervention's true impact on $\hat{\mu}^*_\gamma$. Second, the correlational literature has underestimated the extent to which beliefs matter for parental investment decisions.

A further layer of ambiguity concerns the reference prior itself. The discussion above treats $\mu_{\gamma,0}$ as the benchmark against which $\hat{\mu}^*_\gamma$ is evaluated. But the model takes $\mu_{\gamma,0}$ as given without requiring it to equal the true γ. Two interpretations are possible. Under the first, γ varies across families and $\mu_{\gamma,0}$ accurately reflects the family-specific marginal productivity of responsive investment, so some parents face genuinely lower returns to responsive engagement, for example because of differences in child characteristics or in the complementarity between parental responsiveness and other inputs. Under the second, $\mu_{\gamma,0}$ is itself a biased prior, systematically underestimating γ due to limited exposure to information about child development, cultural norms that discourage verbal interaction with infants, or other forces that shape the reference model before motivated reasoning operates. The two interpretations have different implications for the welfare analysis of the intervention: under the first, raising $\mu_{\gamma,0}$ corrects motivated reasoning about a true return that was already known; under the second, it also corrects a prior bias that predates motivated reasoning. The model's positive predictions about the self-sealing trap, the four escape channels, and the testable implications are identical under both interpretations. But the welfare implications differ: if $\mu_{\gamma,0}$ is already accurate, the intervention restores beliefs to their true level; if $\mu_{\gamma,0}$ is biased, the intervention may raise beliefs above $\mu_{\gamma,0}$ and closer to the true γ, producing welfare gains that exceed what the model attributes to the reduction in motivated reasoning alone. The elicitation procedures used in this paper and in the broader literature measure something between $\mu_{\gamma,0}$ and $\hat{\mu}^*_\gamma$, as argued above, and cannot separately identify whether $\mu_{\gamma,0}$ itself is biased. Disentangling these two sources of belief error (prior bias and motivated distortion) remains an open question for the belief elicitation literature.

Responsive and Non-Responsive Investment The model identifies a further source of downward bias in the estimated belief-investment correlation that has not previously been recognized. In the model, parental beliefs $\hat{\mu}^*_\gamma$ affect the calibrated component of investment (the part that varies with the child's cue in the interior regime, generating $\mathrm{Cov}(x, \theta) > 0$), but not the level component x^R, which is independent of beliefs in both the interior and corner regimes. The true relationship between beliefs and investment is therefore concentrated in the responsive component. The empirical literature, however, measures total investment, as captured, for example, by the Home Observation for the Measurement of the Environment (HOME; Caldwell and Bradley 1981), or by the Family Care Indicators (FCI; Hamadani et al. 2010), which not only aggregates the responsive and non-responsive components, but may have only a few items to measure the responsive component. The non-responsive component is driven by habits, culture, and repertoire rather than beliefs, and including it in the investment measure adds noise that is orthogonal to beliefs, attenuating the estimated belief-investment correlation. This bias is distinct from and cumulative with the social desirability bias documented in the preceding paragraph: one operates on the belief side of the correlation, the other on the investment side, and both push the estimated relationship below the true one. The LENA system's Advanced Data Extractor provides a scalable naturalistic measure of the responsive component. The concentration of treatment effects in child-initiated blocks documented in Section 6 is therefore not only evidence that the intervention shifted the nature of parental investment from uniform to responsive, it is also the most scalable available test of the relationship between beliefs and the component of investment that beliefs actually govern.

8 Conclusion

This paper develops a model of parental investment in early childhood in which bandwidth-constrained parents hold distorted beliefs about the returns to responsive interaction, and tests the model's predictions in a randomized controlled trial. The model identifies believing and acting as two barriers to responsive parenting and shows that low bandwidth makes both barriers harder simultaneously, producing a self-sealing trap in which distorted beliefs suppress engagement, uniform responses generate no calibrating evidence, and the distortion is confirmed. The model further identifies four channels through which interventions can escape the trap, provides a unified language for characterizing what any parent-directed program is doing, and generates testable predictions about what it will and will not achieve. All four predictions are confirmed in the data.

The self-sealing trap characterized in this paper is not specific to parenting. It arises whenever three conditions hold simultaneously: the agent holds accurate beliefs about the returns to a costly action, bandwidth constraints prevent her from fully acting on those beliefs, and uniform behavior eliminates the experiential variation that would challenge the distortion. These conditions are not unique to the parenting context. They arise naturally in any setting where accurate beliefs are psychologically costly because the agent cannot deliver on them, and where the absence of effortful engagement removes the feedback that would sustain accurate beliefs. The model therefore provides a general framework for understanding persistent underinvestment in high-return activities among constrained agents. This phenomenon extends beyond parenting to any domain in which agents must make decisions that require attention, such as health behaviors, human capital investment, financial decision-making, and in which bandwidth constraints, whether from poverty, stress, competing demands, or other sources, create a gap between what accurate beliefs demand and what the agent can deliver.

The four-channel taxonomy developed in this paper provides a design language for parent-directed programs that is independent of the specific application. Any program can be characterized by which channels it activates and the model generates predictions about what the program will and will not achieve before an RCT is run. For example, the LENA Start does not directly address the competing cognitive demands, such as financial stress, irregular work schedules, household management, and others, that deplete the parental bandwidth available for responsive engagement. Interventions that combine direct bandwidth support with belief and cue-detection expertise components would, under the model, produce larger and more persistent effects than either program alone. This remains an open frontier for program design.

A secondary contribution of this paper is measurement. The developmental return to parental investment in the model depends on the covariance between the parent's response and the child's cue ($\mathrm{Cov}(x, \theta) > 0$ in the interior regime), not on the level of adult speech alone. The LENA system's Advanced Data Extractor provides such measure in a scalable fashion: by decomposing conversational turns by initiator, it separates child-initiated exchanges, which is the behavioral signature of a parent reading and responding to child cues, from adult-initiated speech, which reflects investment regardless of the child's signal. The concentration of treatment effects in child-initiated blocks documented in Section 6 is direct evidence that the intervention shifted the nature of parental investment from uniform to responsive, not merely its level. Whether equivalent scalable measurement infrastructure can be developed for other developmental domains where responsive parenting also matters, remains an open question for the measurement literature.

The randomized controlled trial reported in this paper measures effects over a ten-week horizon. Whether the belief and engagement effects documented here translate into durable developmental gains for the children in the sample remains an open question. A follow-up study will assess children's language and cognitive development at kindergarten entry (about three years after entering the LENA Start Program). The model generates a specific prediction for this follow-up: children whose parents were in the treatment group should show larger developmental gains than children whose parents were in the control group. Whether the data will support this prediction is a question we leave to future work.

References

Ainsworth, Mary D. S. (1985). "Attachments across the life span". In: *Bulletin of the New York Academy of Medicine* 61.9, pp. 792–812.

Ainsworth, Mary D. S. et al. (1978). *Patterns of Attachment: A Psychological Study of the Strange Situation*. Hillsdale, NJ: Lawrence Erlbaum Associates. DOI: 10.1037/t28248-000.

Akabayashi, Hideo (2006). "An equilibrium model of child maltreatment". In: *Journal of Economic Dynamics & Control* 30, pp. 993–1025.

Almond, Douglas and Janet Currie (2011). "Human Capital Development Before Age Five". In: *Handbook of Labor Economics*. Ed. by O. Ashenfelter and D. Card. Vol. 4. Elsevier, pp. 1315–1486.

Almond, Douglas, Janet Currie, and Valentina Duque (2018). "Childhood Circumstances and Adult Outcomes: Act II". In: *Journal of Economic Literature* 56.4, pp. 1360–1446.

Athey, Susan, Julie Tibshirani, and Stefan Wager (Apr. 2019). "Generalized Random Forests". In: *The Annals of Statistics* 47.2, pp. 1148–1178. ISSN: 0090-5364, 2168-8966. DOI: 10.1214/18-AOS1709. https://projecteuclid.org/journals/annals-of-statistics/volume-47/issue-2/Generalized-random-forests/10.1214/18-AOS1709.full.

Attanasio, Orazio, Flávio Cunha, and Pamela Jervis (Dec. 2019). *Subjective Parental Beliefs: Their Measurement and Role*. Working Paper 26516. National Bureau of Economic Research. DOI: 10.3386/w26516. https://www.nber.org/papers/w26516.

Attanasio, Orazio P. (2026). "The first 1000 days and beyond: The process of child development". In: *Journal of Economic Literature*. Forthcoming.

Attanasio, Orazio P. et al. (2014). "Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial". In: *BMJ* 349, g5785.

Bénabou, Roland (2015). "The Economics of Motivated Beliefs". In: *Revue d'Économie Politique* 125.5, pp. 665–685. DOI: 10.3917/redp.255.0665. https://www.cairn.info/revue-d-economie-politique-2015-5-page-665.htm.

Bénabou, Roland and Jean Tirole (Aug. 2002). "Self-Confidence and Personal Motivation". In: *Quarterly Journal of Economics* 117.3, pp. 871–915. DOI: 10.1162/003355302760193913. https://doi.org/10.1162/003355302760193913.

Bhalotra, Sonia et al. (Apr. 2025). "Maternal Investments in Children: The Role of Expected Effort and Returns". In: *The Economic Journal* 135.667, pp. 712–747. DOI: 10.1093/ej/ueae059. https://doi.org/10.1093/ej/ueae059.

Boneva, Teodora and Christopher Rauh (Dec. 2018). "Parental Beliefs about Returns to Educational Investments—The Later the Better?" In: *Journal of the European Economic Association* 16.6, pp. 1669–1711. DOI: 10.1093/jeea/jvy006. https://doi.org/10.1093/jeea/jvy006.

Brunnermeier, Markus K. and Jonathan A. Parker (Sept. 2005). "Optimal Expectations". In: *American Economic Review* 95.4, pp. 1092–1118. DOI: 10.1257/0002828054845280. https://www.aeaweb.org/articles?id=10.1257/0002828054845280.

Caldwell, Bettye M. and Robert H. Bradley (1981). "The home observation for measurement of the environment: psychometric properties and uses". In: *Developmental Review* 1.2, pp. 113–145.

Caplin, Andrew, Mark Dean, and John Leahy (Nov. 2022). "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy". In: *Econometrica* 90.6, pp. 2471–2501. DOI: 10.3982/ECTA16081. https://doi.org/10.3982/ECTA16081.

Castagnetti, Alessandro and Renke Schmacker (2022). "Protecting the ego: Motivated information selection and updating". In: *European Economic Review* 142, p. 104007. ISSN: 0014-2921. DOI: https:

//doi.org/10.1016/j.euroecorev.2021.104007. https://www.sciencedirect.com/science/article/pii/S0014292121002786.

Cates, Carolyn Brockmeyer et al. (2023). "Validation of the StimQ2: A parent-report measure of cognitive stimulation in the home". In: *PLoS ONE* 18.7, e0286708. DOI: 10.1371/journal.pone.0286708.

Celik-Katreniak, Dagmara (July 2021). "Overconfidence for happiness? Evidence from Uganda". SSRN Working Paper 3284473. DOI: 10.2139/ssrn.3284473. https://ssrn.com/abstract=3284473.

Cheremukhin, Anton, Anna Popova, and Antonella Tutino (2011). "Experimental evidence on rational inattention". In: https://api.semanticscholar.org/CorpusID:73700442.

Chetty, Raj and Adam Szeidl (2016). "Consumption Commitments and Habit Formation". In: *Econometrica* 84.2, pp. 855–890. DOI: https://doi.org/10.3982/ECTA9390. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA9390. https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA9390.

Cobb-Clark, Deborah A., Nicolás Salamanca, and Anna Zhu (2019). "Parenting Style as an Investment in Human Development". In: *Journal of Population Economics* 32.4, pp. 1315–1352. DOI: 10.1007/s00148-018-0703-2.

Conti, Gabriella, Michele Giannola, and Alessandro Toppeta (2022). *Parental Beliefs, Perceived Health Risks, and Time Investment in Children: Evidence from COVID-19*. Working Paper 658. Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy. https://www.csef.it/WP/wp658.pdf.

Cosconati, Marco (2008). "Parenting Style and the Development of Human Capital in Children". PhD thesis. Philadelphia, PA: The University of Pennsylvania.

Cristia, Alejandrina, Federica Bulgarelli, and Elika Bergelson (2020). "Accuracy of the Language Environment Analysis System Segmentation and Metrics: A Systematic Review". In: *Journal of Speech, Language, and Hearing Research* 63.4, pp. 1093–1105. DOI: 10.1044/2019_JSLHR-19-00386.

Cunha, Flavio and James Heckman (2007). "The technology of skill formation". In: *American Economic Review* 97.2, pp. 31–47.

Cunha, Flavio et al. (Mar. 2025). *An Evaluation of the Alief Independent School District Jumpstart Program: Using a Model to Recover Mechanisms from an RCT*. Working Paper 33537. National Bureau of Economic Research.

Cunha, Flávio, Irma Elo, and Jennifer Culhane (June 2013). *Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation*. w19144. Cambridge, MA: National Bureau of Economic Research, w19144. DOI: 10.3386/w19144. http://www.nber.org/papers/w19144.pdf.

– (Nov. 1, 2022). "Maternal Subjective Expectations about the Technology of Skill Formation Predict Investments in Children One Year Later". In: *Journal of Econometrics*. Annals Issue: Subjective Expectations & Probabilities in Economics 231.1, pp. 3–32. ISSN: 0304-4076. DOI: 10.

1016/j.jeconom.2020.07.044. https://www.sciencedirect.com/science/article/pii/S0304407620302700.

Cunha, Flávio et al. (Mar. 2024). "Language Environment and Maternal Expectations: An Evaluation of the LENA Start Program". In: *Journal of Human Capital* 18.1, pp. 105–139. ISSN: 1932-8575. DOI: 10.1086/729038. https://www.journals.uchicago.edu/doi/abs/10.1086/729038.

Del Boca, Daniela et al. (2026). "Parenting with Patience: Parental Incentives and Child Development". In: *Journal of Political Economy* 134.1.

Doepke, Matthias and Fabrizio Zilibotti (Sept. 2017). "Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission". In: *Econometrica* 85.5, pp. 1331–1371.

Drobner, Christoph and Sebastian J. Goerg (2024). "Motivated Belief Updating and Rationalization of Information". In: *Management Science* 70.7, pp. 4583–4592. DOI: 10.1287/mnsc.2023.02537.

Dupas, Pascaline et al. (2025). "Informing Mothers about the Benefits of Conversing with Infants: Experimental Evidence from Ghana". In: *American Economic Journal: Economic Policy* 17.2, pp. 388–417. DOI: 10.1257/pol.20230283.

Falk, Armin et al. (Nov. 1, 2018). "Global Evidence on Economic Preferences*". In: *The Quarterly Journal of Economics* 133.4, pp. 1645–1692. ISSN: 0033-5533. DOI: 10.1093/qje/qjy013. https://doi.org/10.1093/qje/qjy013.

Fisher, Philip A. et al. (2016). "Promoting Healthy Child Development via a Two-Generation Translational Neuroscience Framework: The Filming Interactions to Nurture Development Video Coaching Program". In: *Child Development Perspectives* 10.4, pp. 251–256. DOI: 10.1111/cdep.12195.

Geeraerts, Sanne B. et al. (2021). "The Role of Parental Self-Regulation and Household Chaos in Parent-Toddler Interactions: A Time-Series Study". In: *Journal of Family Psychology* 35.2, pp. 236–246. DOI: 10.1037/fam0000814.

Gertler, Paul et al. (2014). "Labor market returns to an early childhood stimulation intervention in Jamaica". In: *Science* 344.6187, pp. 998–1001. DOI: 10.1126/science.1130121.

Gertler, Paul et al. (2021). *Effect of the Jamaica Early Childhood Stimulation Intervention on Labor Market Outcomes at Age 31*. Working Paper 29292. National Bureau of Economic Research. DOI: 10.3386/w29292.

Giannola, Michele (2023). "Parental Investments and Intra-household Inequality in Child Human Capital: Evidence from a Survey Experiment". In: *The Economic Journal* 134.658, pp. 671–727. DOI: 10.1093/ej/uead086.

Gilkerson, Jill and Jeffrey A. Richards (July 2020). *A Guide to Understanding the Design and Purpose of the LENA System*. Tech. rep. LTR-12. Boulder, CO: LENA Foundation.

Grantham-McGregor, Sally M, Susan P Powell Christine A Walker, and John H Himes (1991). "Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican Study". In: *The Lancet* 338.8758, pp. 1–5. DOI: 10.1016/0140-6736(91)90001-6.

Hagger, M. S. et al. (2016). "A Multilab Preregistered Replication of the Ego-Depletion Effect". In: *Perspectives on Psychological Science* 11.4, pp. 546–573. DOI: 10.1177/1745691616652873.

Hamadani, Jena D. et al. (2010). "Use of family care indicators and their relationship with child development in Bangladesh". In: *Journal of Health, Population, and Nutrition* 28.1, pp. 23–33. DOI: 10.3329/jhpn.v28i1.4520.

Hanell, Tomas (2022). "Unmet Aspirations and Urban Malaise". In: *Social Indicators Research* 164, pp. 83–103. DOI: 10.1007/s11205-021-02864-8.

Heckman, James J. and Stefano Mosso (2014). "The Economics of Human Development and Social Mobility". In: *Annual Review of Economics* 6.1, pp. 689–733. DOI: 10.1146/annurev-economics-080213-040753.

Heckman, James J. et al. (2017). *An Analysis of the Memphis Nurse-Family Partnership Program*. Working Paper 23610. National Bureau of Economic Research. DOI: 10.3386/w23610.

Hyland, David and Mahault Albarracin (2025). *On the Variational Costs of Changing Our Minds*. arXiv: 2509.17957 [cs.AI]. https://arxiv.org/abs/2509.17957.

Kenny, Méabh and Oliver C. Robinson (2026). "Investigating Relationships Between Guilt, Shame, Self-discrepancy and Fear of Negative Evaluation in Mothers: A Replication and Extension". In: *Journal of Child and Family Studies*. DOI: 10.1007/s10826-026-03274-5.

Kőszegi, Botond (2006). "Ego Utility, Overconfidence, and Task Choice". In: *Journal of the European Economic Association* 4.4, pp. 673–707. DOI: 10.1162/JEEA.2006.4.4.673. https://doi.org/10.1162/JEEA.2006.4.4.673.

Landry, Susan H., Karen E. Smith, and Paul R. Swank (2006). "Responsive parenting: Establishing early foundations for social, communication, and independent problem-solving skills." In: *Developmental Psychology* 42.4, pp. 627–642. DOI: 10.1037/0012-1649.42.4.627.

Lavechin, Marvin et al. (2020). "An open-source voice type classifier for child-centered daylong recordings". In: *arXiv preprint arXiv:2005.12656*. https://arxiv.org/abs/2005.12656.

Lee, David S. (2009). "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects". In: *Review of Economic Studies* 76.3, pp. 1071–1102. DOI: 10.1111/j.1467-937x.2009.00536.x.

Lim, Sue Siang (2020). "The Effects of Chronic Illness on Aspirations and Subjective Wellbeing". In: *Journal of Happiness Studies* 21, pp. 1771–1793. DOI: 10.1007/s10902-019-00156-0.

Liss, Miriam, Holly H. Schiffrin, and Kristin M. Rizzo (2013). "Maternal Guilt and Shame: The Role of Self-Discrepancy and Fear of Negative Evaluation". In: *Journal of Child and Family Studies* 22, pp. 1112–1119. DOI: 10.1007/s10826-012-9673-2.

List, John A., Julie Pernaudet, and Dana L. Suskind (2021). "Shifting parental beliefs about child development to foster parental investments and improve school readiness outcomes". In: *Nature Communications* 12.1, p. 5765. DOI: 10.1038/s41467-021-25964-y.

Lizzeri, Alessandro and Marciano Siniscalchi (2008). "Parental Guidance and Supervised Learning". In: *The Quarterly Journal of Economics* 123.3, pp. 1161–1195. DOI: 10.1162/qjec.2008.123.3.1161.

Luchkina, Elena and Fei Xu (2024). "How does social contingency facilitate vocabulary development?" In: *Developmental Science* 27.5, e13525. DOI: 10.1111/desc.13525. https://doi.org/10.1111/desc.13525.

Mackowiak, Bartosz, Filip Matejka, and Mirko Wiederholt (2018). "Dynamic rational inattention: Analytical results". In: *J. Econ. Theory* 176, pp. 650–692. https://api.semanticscholar.org/CorpusID:49695381.

Mani, Anandi et al. (2013). "Poverty Impedes Cognitive Function". In: *Science* 341.6149, pp. 976–980. DOI: 10.1126/science.1238041. https://doi.org/10.1126/science.1238041.

Matějka, Filip and Alisdair McKay (Jan. 2015). "Rational Inattention to Discrete Choices: A New Foundation for Generalized Logit Models". In: *American Economic Review* 105.1, pp. 272–298. DOI: 10.1257/aer.20130047. https://www.aeaweb.org/articles?id=10.1257/aer.20130047.

McGillion, Michelle L. et al. (2013). "Supporting Early Vocabulary Development: What Sort of Responsiveness Matters?" In: *IEEE Transactions on Autonomous Mental Development* 5, pp. 240–248. https://api.semanticscholar.org/CorpusID:2681253.

Oh, Hyunjoo, Dong-gwi Lee, and Hunggu Cho (2023). "The differential roles of shame and guilt in the relationship between self-discrepancy and psychological maladjustment". In: *Frontiers in Psychology* 14, p. 1215177. DOI: 10.3389/fpsyg.2023.1215177.

Olds, David L. et al. (1986). "Improving the delivery of prenatal-care and outcomes of pregnancy—A randomized trial of nurse home visitation". In: *Pediatrics* 77.1, pp. 16–28.

Roggman, Lori A. et al. (2013). "Parenting Interactions with Children: Checklist of Observations Linked to Outcomes (PICCOLO) in diverse ethnic groups". In: *Infant Mental Health Journal* 34.5, pp. 1–14. DOI: 10.1002/imhj.21389.

Romeo, Rachel R. et al. (May 2018). "Beyond the 30-Million-Word Gap: Children's Conversational Exposure Is Associated With Language-Related Brain Function". In: *Psychological Science* 29.5, pp. 700–710. DOI: 10.1177/0956797617742725. https://doi.org/10.1177/0956797617742725.

Seror, Avner (2022). "Child Development in Parent-Child Interactions". In: *Journal of Political Economy* 130.9, pp. 2462–2499. DOI: 10.1086/720398.

Shah, Anuj K., Sendhil Mullainathan, and Eldar Shafir (2012). "Some Consequences of Having Too Little". In: *Science* 338.6107, pp. 682–685. DOI: 10.1126/science.1222426.

Sims, Christopher A. (Apr. 2003). "Implications of Rational Inattention". In: *Journal of Monetary Economics* 50.3, pp. 665–690. DOI: 10.1016/S0304-3932(03)00029-1. https://doi.org/10.1016/S0304-3932(03)00029-1.

Slobodin, Ortal et al. (2020). "Mothers' Need Frustration and Controlling Parenting: The Moderating Role of Maternal Guilt". In: *Journal of Child and Family Studies* 29, pp. 1914–1926. DOI: 10.1007/s10826-020-01720-6.

Stone, Jeff and Joel Cooper (2001). "A Self-Standards Model of Cognitive Dissonance". In.

Suskind, Dana L. et al. (Mar. 2016). "A Parent-Directed Language Intervention for Children of Low Socioeconomic Status: A Randomized Controlled Pilot Study". In: *Journal of Child Language* 43.2, pp. 366–406. ISSN: 0305-0009, 1469-7602. DOI: 10.1017/S0305000915000033. https://www.cambridge.org/core/journals/journal-of-child-language/article/abs/parentdirected-language-intervention-for-children-of-low-socioeconomic-status-a-randomized-controlled-pilot-study/C6707146A62899FA2834C85D77997418.

Tamis-LeMonda, Catherine S., Yana Kuchirko, and Lulu Song (2014). "Why Is Infant Language Learning Facilitated by Parental Responsiveness?" In: *Current Directions in Psychological Science* 23, pp. 121–126. https://api.semanticscholar.org/CorpusID:43289662.

Tate, A. D. et al. (2024). "Association between parental resource depletion and parent use of specific food parenting practices: An Ecological Momentary Assessment Study". In: *Appetite* 199, p. 107368. DOI: 10.1016/j.appet.2024.107368.

Thomas, Gregory Owen, Wouter Poortinga, and Elena Sautkina (2016). "Habit Discontinuity, Self-Activation, and the Diminishing Influence of Context Change: Evidence from the UK Understanding Society Survey". In: *PLoS ONE* 11.4, e0153490. DOI: 10.1371/journal.pone.0153490.

Thunström, Linda et al. (Apr. 2016). "Strategic Self-Ignorance". In: *Journal of Risk and Uncertainty* 52.2, pp. 117–136. DOI: 10.1007/s11166-016-9236-9.

Verriest, Ewout (2024). "Household Labor Supply, Child Development and Childcare Policies in the United States". Unpublished manuscript, April 8, 2024, Pennsylvania State University.

Walker, Sarah et al. (2024). *HeForShe: Bargaining Power, Parental Beliefs and Parental Investments*. CEPR Discussion Paper 18775. Centre for Economic Policy Research. https://cepr.org/publications/dp18775.

Wood, Wendy, Leona Tam, and Melissa Guerrero Witt (2005). "Changing Circumstances, Disrupting Habits". In: *Journal of Personality and Social Psychology* 88.6, pp. 918–933. DOI: 10.1037/0022-3514.88.6.918.

Young, Alwyn (May 1, 2019). "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results*". In: *The Quarterly Journal of Economics* 134.2, pp. 557–598. ISSN: 0033-5533. DOI: 10.1093/qje/qjy029. https://doi.org/10.1093/qje/qjy029.

Zhang, Xutong et al. (2022). "A dynamic systems account of parental self-regulation processes in the context of challenging child behavior". In: *Child Development* 93.5, e501–e514. DOI: 10.1111/cdev.13808.