# From Proto-Agent to Person: A Bootstrapping Architecture for the Development of Personhood

PIATRA . INSTITUTE

June 2026

## Abstract

A human newborn is a living organism with homeostatic regulation, affect, sensorimotor coupling, social-orienting biases, and a powerful capacity for statistical learning. It is not yet a language-bearing, norm-bearing, autobiographically extended self. By the age of three to five a recognizable someone has formed, and by adulthood the same initial architecture can support a scientist, a citizen, a parent, a legal subject, a reflective philosophical agent. This paper proposes a bootstrapping architecture that accounts for that divergence. It takes as its starting point the interactionist consensus already established by neuroconstructivism and developmental systems theory, that development is neither maturation alone nor inscription on a blank slate but emerges from organism-environment coupling, and its contribution is not that premise but a specific staging of it. The thesis is that adult cognitive personhood is co-compiled: biologically prepared proto-agential systems enter caregiver-mediated loops of affect regulation, joint attention, language, norm enforcement, autobiographical narration, executive scaffolding, and cultural-tool use, and the child progressively internalizes the regulatory and symbolic structures first run around it by others. The argument is an integrative synthesis with a defined scope. It stages a determinate proto-agent-to-person stack and integrates developmental psychology, attachment theory, predictive processing, language-acquisition research, Vygotskian internalization, Tomasellian shared intentionality, autobiographical-memory theory, executive-function development, and cultural evolution, carrying the bootstrapping process past childhood into the institutional layers that most developmental syntheses leave unaddressed, and it tests the integrated account against the boundary cases of institutional deprivation and late first-language acquisition. The result is a layered control architecture in which each acquired layer rewrites the learning conditions for the next. The paper is candid that an integration of this breadth is hard to falsify as a whole and that much of what it accommodates it accommodates by construction, and it isolates the mediation and ordering claims on which the architecture could actually be tested. The account is developmental and cognitive, and is held strictly separate from claims about moral or legal status: the newborn lacks adult-style cognitive personhood and retains full moral standing. The developmental layering mirrors, under compression, the layered history of the person concept itself, which classical anthropology has long treated as a category assembled across the history of societies rather than given whole.

## 1. The bootstrapping problem

The same newborn brain can become a non-linguistic survivor of extreme deprivation, an ordinary social adult, or a research mathematician who manipulates objects no perceptual system evolved to track. The range is enormous and the starting hardware is, to a first approximation, constant. Whatever explains the spread is not encoded in the genome in any direct sense, because the genome is shared, and it is not supplied whole by the environment either, because a tape recorder left running in an empty room teaches an infant almost nothing. The explanation has to be a developmental process that couples a specific kind of organism to a specific kind of surround and produces, over time, capacities that were present in neither alone.

Two familiar answers fail. Strong nativism holds that personhood unfolds from an internal pro-

gram given biological time and adequate nutrition. It cannot explain why children deprived of accessible language in the first years never fully acquire grammar afterward, or why institutionally reared infants show lasting deficits in memory and executive control that nutrition does not repair. Strong environmentalism holds that culture writes personhood onto a passive surface. It cannot explain why the infant is active from the first day, recruiting caregivers, predicting contingencies, orienting to faces and voices, resisting and exploring, nor why temperament, sensitive periods, and embodied constraints shape what any environment can install.

The account defended here is neither. Personhood is bootstrapped. A small set of biologically prepared capacities is sufficient to enter a structured social environment as a participant. Participation generates new learning data. That data supports the acquisition of a new capacity, which changes how the child participates, which changes the data again. The loop is recursive, and the recursion is the point. Language changes what can be remembered. Memory changes what counts as the self. The self changes what is worth doing. Norms change how actions are selected. Schooling changes what attention is for. At each turn the system that does the next round of learning is not the system that did the last.

It must be said plainly that "neither nativism nor environmentalism" is not a new position. It is the standing consensus of developmental science, and the better part of a research tradition has already been built on it. Karmiloff-Smith's representational-redescription account replaced fixed innate modules with a process by which the developing mind progressively rerepresents its own implicit knowledge into explicit, flexible forms, dissolving the nativist-empiricist dichotomy at the level of cognitive architecture (Karmiloff-Smith, 1992). The neuroconstructivist programme of Mareschal and colleagues makes the same move at the level of brain development, treating representations as emergent products of a brain whose own structure is shaped by its activity in an environment, so that gene, neuron, body, and context are partial constraints on one developmental trajectory rather than competing sources of a finished design (Mareschal et al., 2007). Developmental systems theory states the most radical version: Oyama argues that information is not transmitted ready-made by genes or imposed ready-made by the environment but constructed anew in each life cycle by the whole developmental system, and Griffiths and Gray generalize the point into an account of inheritance and evolution in which no single factor is privileged as the carrier of form (Oyama, 1985; Griffiths & Gray, 1994). The interactionist conclusion that development emerges from organism-environment coupling rather than from maturation or instruction alone is, in these literatures, settled. No serious developmentalist defends the strong nativism or strong environmentalism that the two-failed-answers framing rejects, and the paper does not claim to refute live opponents by rejecting them.

What this paper adds is therefore not the interactionist thesis but a specific use of it. The neuroconstructivist and developmental-systems literatures establish that development is an interaction and analyze its mechanics in depth, but they do not stage the route from the neonate to the institutionally competent adult as a determinate sequence of named layers, and they say comparatively little about the upper end of that route, the symbolic, normative, narrative, and institutional capacities installed across adolescence and adulthood. The contribution here is twofold. First, it

commits to a specific staging, the proto-agent-to-person stack of section 2, in which each layer rewrites the learning conditions for the next and the direction of travel is the inward migration of control. Second, it integrates literatures that are usually kept apart, co-regulation and attachment, joint attention and shared intentionality, language as symbolic control, norm internalization, autobiographical memory, executive function, and cumulative culture, treating them not as separate research topics but as successive phases of one bootstrapping process, and it carries that process past childhood into the institutional layers that most developmental syntheses leave unaddressed. The closest mainstream anchors treat fragments of this loop. Thompson's synthesis of social understanding, relationships, self, and conscience is the nearest mainstream developmental account of "the development of the person" as such (Thompson, 2006). Tomasello's theory of ontogeny derives uniquely human cognition from shared intentionality and cultural learning (Tomasello, 2019). Nelson and Fivush model autobiographical memory as a social-cultural achievement assembled from language, adult memory talk, and temporal understanding (Nelson & Fivush, 2004). What none supplies is the full vertical integration: a single staged architecture in which co-regulation, language, norms, narrative, executive control, and cumulative culture are successive layers of one process running from the neonate to the institution-compatible adult. This paper supplies that integration. Its contribution is synthetic rather than the discovery of a new mechanism, and the rest of the paper is honest about that scope.

## 2. Conceptual distinctions

The word "person" holds at least three distinct meanings that the rest of the argument must keep apart. There is the *moral* person, the bearer of intrinsic worth and rights, which a newborn possesses in full. There is the *legal* person, the bearer of formal standing within an institutional order, which Roman law constructed as a status that can be assigned, expanded, or withdrawn. And there is the *developmental-cognitive* person, the organized capacity for self-regulation, symbolic thought, accountability, and temporally extended identity, which is acquired over the course of childhood and adolescence. This paper concerns the third sense only. The newborn lacks adult-style cognitive personhood and possesses undiminished moral standing. Holding these apart prevents the single most common misreading of the developmental claim, which is to hear a statement about cognitive assembly as a statement about who deserves protection.

Within the cognitive sense, personhood is not one capacity but a stack of layers. The layers overlap, develop unevenly, regress under stress, and vary across cultures, and the broad direction of travel is constant: control migrates inward. The caregiver first regulates the infant from outside; the mature agent regulates itself through internalized models, language, norms, memory, and role expectations.

| Layer | Minimal description | What it adds relative to the layer below |
|---|---|---|
| Organism | Living human body with metabolism, homeostasis, sleep, feeding, arousal, reflexive regulation | Biological viability, present before any agency |
| Proto-agent | Affective, embodied system that orients, seeks comfort, avoids distress, learns contingencies | Action becomes organized around needs, though not yet symbolic or self-owned |
| Embodied agent | Sensorimotor learner with body schema, action prediction, object tracking, locomotion | The world becomes a field of manipulable possibilities |
| Social agent | Participant in dyadic exchange: gaze, affect, turn-taking, attachment, joint attention | Other minds become regulators, interpreters, and targets of action |
| Language-using self | User of names, categories, pronouns, requests, negation, tense, and eventually inner speech | Experience becomes symbolically structured and publicly shareable |
| Narrative self | Agent who links events into a remembered, owned, temporally ordered first-person sequence | Memory becomes autobiographical rather than merely episodic or procedural |
| Moral person | Agent answerable to rules, obligations, fairness, blame, praise, promises, repair | Behavior becomes conduct, interpretable through shared standards |
| Institution-compatible adult | Agent legible to schools, law, medicine, money, documents, professions, citizenship | Identity becomes addressable by bureaucratic, legal, and economic systems |
| Reflective self | Agent who models its own beliefs, motives, mortality, ideals, and place in abstract systems | The self becomes an object of explicit inquiry and revision |

The minimal-self literature gives the lowest two transitions a useful sharpening. Gallagher distinguishes a *minimal self*, the immediate embodied for-me-ness of experience, from a *narrative self*, the temporally extended story of who one is (Gallagher, 2000). Infants plausibly have something like

the first long before they have anything like the second. The developmental question is how the second is built on top of the first, and the answer occupies the rest of this paper.

## 3. The newborn as embodied proto-agent

The newborn is neither a blank slate nor a miniature adult. It is a highly plastic, socially dependent, homeostatic prediction-and-control system, and the temptation to over-read its competence is as distorting as the temptation to under-read it.

What is present early is a biological control architecture. The infant breathes, feeds, cycles through sleep and wakefulness, and signals distress, though much of this stability is supplied from outside by caregivers who regulate temperature, satiation, and arousal. Affect is present in the form of distress, comfort, interest, satiation, and startle, though not in the form of guilt, shame, pride, or resentment, which require concepts the infant does not yet hold. Sensorimotor control exists in primitive form: reflexes, sucking, orienting, early gaze, grasping. Caregiver-seeking is robust from the start, in crying, calming to touch and voice, and proximity regulation, and it is best read as the infant recruiting adults as external regulators rather than as mere reflex discharge.

Two early capacities do most of the bootstrapping work. The first is a set of social-orienting biases. Newborns preferentially track face-like configurations (Johnson et al., 1991), prefer infant-directed speech (Cooper & Aslin, 1990), prefer speech to closely matched non-speech (Vouloumanos & Werker, 2007), discriminate their native language by its rhythm within days of birth (Mehler et al., 1988), and orient to biological motion (Simion, Regolin & Bulf, 2008). These biases do not amount to social cognition. They amount to a sampling policy that points the plastic learner at the stimuli that bear social structure, other people. The second is statistical learning. Eight-month-olds extract regularities from the speech stream, tracking the transitional probabilities between syllables well enough to segment a continuous artificial language into word-like units after about two minutes of exposure (Saffran, Aslin & Newport, 1996). This is the substrate on which phonology, word segmentation, and pattern learning are later built.

Predictive processing offers a unifying frame for these scattered abilities. On the free-energy formulation, an agent acts to minimize the long-run prediction error, or surprise, of its sensory states (Friston, 2010; Clark, 2013), learning contingencies through perception-and-action loops: crying brings the caregiver, sucking brings milk, a turn of the head brings a face into view. Köster and colleagues argue that statistical learning, proprioceptive and motor learning, and early physical and social understanding can be read as instances of one prediction-error-minimizing process operating over experience-expectant priors (Köster et al., 2020). The frame is useful because it makes the infant active. The infant samples, predicts, and acts to bring its predictions about a deeply social world into register, from the first days.

Several claims about newborns should be made carefully. Neonatal imitation, long treated as evidence of an innate matching system, is contested: a large longitudinal study failed to find reliable imitation of modeled gestures across the first nine weeks (Oostenbroek et al., 2016), a null that has itself been disputed (Meltzoff et al., 2018), and a meta-analysis of 336 effect sizes recovered a

positive pooled estimate alongside substantial unexplained heterogeneity (Davis et al., 2021). It is safer to speak of early sensorimotor resonance and a readiness for social learning than of adult-like imitation at birth. Explicit theory of mind, measured as above-chance performance on standard false-belief tasks, consolidates around the fourth year (Wellman, Cross & Watson, 2001), although infants show earlier implicit sensitivity to others' false beliefs in looking-time paradigms (Onishi & Baillargeon, 2005). Visual self-recognition, the mark of an objective self that can become an object to itself, emerges in the second year (Lewis & Brooks-Gunn, 1979; Rochat, 2003). Stable autobiographical memory, adult grammar, explicit norms, and institutional agency are absent in any mature form. The newborn architecture is an embodied biological control system with affective priorities, social-orienting biases, plastic prediction-driven learning, and a strong structural expectation of external regulation.

## 4. Co-regulation before self-regulation

The first thing the social environment supplies is regulation. Before the infant can regulate its own arousal, hunger, fear, or attention, the caregiver does it for them, and the manner in which this is done becomes the infant's first model of how the world responds.

Affect regulation begins as a two-body process. The caregiver feeds, holds, rocks, and soothes, and in doing so brings the infant's physiological and emotional state back inside tolerable bounds. Repeated across thousands of episodes, this co-regulation is not merely palliative. It teaches a generalizable expectation: distress is followed by relief, signals are answered, states that exceed the infant's own capacity can be brought down by another. The regulation is physiological before it is psychological. Caregivers act as external regulators of the infant's autonomic and interoceptive state, and on one account the social brain grows through this allostatic co-regulation, the caregiver anticipating and correcting the infant's internal balance until the infant can do so itself (Atzil et al., 2018). Taipale's analysis makes the developmental logic explicit: mature affective self-regulation develops by internalizing regulatory functions that were initially performed by the caregiver, so that the dyadic pattern of soothing becomes the individual capacity to self-soothe (Taipale, 2016). The infant first learns that regulation happens through this other. Only later does the same pattern run inside one body.

Attachment theory describes what the infant learns about the regulator. Repeated patterns of response build an internal model of whether comfort is available, predictable, intrusive, or absent, and this model organizes the infant's use of the caregiver as a secure base for exploration and a safe haven in distress (Bowlby, 1969; Ainsworth et al., 1978). The model is consequential because exploration depends on it. A child who expects reliable return can leave the caregiver to investigate the world; a child who does not must spend regulatory resources monitoring an unreliable base. Caregiver sensitivity matters here without being destiny: a meta-analysis found a significant but modest association between parental sensitivity and attachment security ($r = .22$ across 30 studies, $N = 1{,}666$; the larger pool of 66 studies covered all parental antecedents combined rather than sensitivity alone), which leaves substantial room for temperament, context, and later relationships (De Wolff & van IJzendoorn, 1997).

The active-inference reading of attachment treats the internal working model as a learned prior over social availability. On this account the infant forms expectations about whether signaling will reduce uncertainty and distress, and the secure, avoidant, and ambivalent patterns correspond to different policies for minimizing free energy over interoceptive stress states when the caregiver's responses are reliable, reliably unresponsive, or inconsistent (Cittern et al., 2018). The reading connects attachment to the same prediction-and-action machinery that drives early learning generally, which is what allows co-regulation to be the first layer of a single architecture rather than a separate emotional subsystem.

A caution belongs here. The external regulator is not always benign. The same loop that compiles trust, recovery, and exploratory confidence can compile vigilance, dysregulation, and the expectation that signals go unanswered. The architecture is neutral about the content of what gets installed. It guarantees only that something will be, because the infant cannot not learn from the regulator it has.

## 5. Social address and the creation of a someone

The infant is treated as a person before it is one in the cognitive sense, and this premature treatment is itself a mechanism. Caregivers name the infant, address it in the second person, attribute intentions and feelings to its movements, and interpret its cries as meaningful communications about a world. The infant cannot yet support these attributions. The attributions run anyway, and they structure the interactions from which the infant learns.

The mechanism is visible in how adults read intention into behavior that does not yet hold it. A reach is treated as wanting, a gaze as looking-at, a cry as asking-for. The caregiver responds to the imputed intention, and the response makes the imputed structure real in the next round: the infant, finding that reaching reliably produces the reached-for object via an adult, learns to reach as a communicative act. The "someone" the caregiver addresses is, at first, a useful fiction. The fiction is productive because acting on it generates the contingencies that eventually make it true.

Affect mirroring is the most studied case. Gergely and Watson's social-biofeedback theory holds that caregivers reflect infant affect in marked, exaggerated, and well-timed displays, and that the infant uses these displays as external feedback about its own internal states (Gergely & Watson, 1996). The marking matters: the caregiver's display is recognizably a representation of the infant's state rather than the caregiver's own, which lets the infant bind a diffuse bodily-affective pattern to a categorized, shareable emotion. On the mentalization account, this marked mirroring is also the route by which a caregiver's reading of the infant's mind scaffolds the infant's own affect regulation and, later, its capacity to represent mental states as such (Fonagy, Gergely, Jurist & Target, 2002). This is the beginning of emotional self-awareness, and it is supplied from outside before it can be generated within.

The functions caregivers perform across these early loops can be read together as a temporary external runtime for personhood. The adult supplies, from outside, regulatory and representational operations the child cannot yet perform, and runs them in the child's presence and on the child's

behalf until the child can emulate them. The table collects the operations that the preceding sections and the ones that follow assign to the caregiver.

| Adult function | Operation supplied from outside | What the child later internalizes |
|---|---|---|
| External regulator | Stabilizing arousal, hunger, fear, attention | Self-regulation |
| Attention director | Pointing, showing, naming the relevant | Endogenous control of attention |
| Symbolic interpreter | Mapping world and body into signs | Symbolic self-representation |
| Norm enforcer | Correction, permission, the error signal | Self-imposed accountability |
| Memory co-author | Storing and reconstructing early events | Autobiographical narration |
| Identity assigner | Name, kinship, category, family narrative | A self-description to revise |
| Cultural package manager | Installing tools, scripts, rituals, texts | Internalized cultural competence |
| Executive prosthesis | Planning, sequencing, inhibiting, reminding | Endogenous executive control |

The runtime description is a heuristic and has a sharp limit. The brain is an embodied, self-organizing system with evolved priors, temperamental variation, and sensitive periods rather than general-purpose hardware idling until an operating system arrives, and the child runs its own processes throughout: predicting, recruiting, imitating, resisting, exploring, and by its responses reshaping the very environment that scaffolds it. The runtime is also indifferent to the content it runs. The same loop that installs trust, curiosity, and self-command can install fear, shame, and learned helplessness, because the child internalizes the regulation it actually receives rather than the regulation it needs. What the heuristic names accurately is a transfer: operations that begin distributed across a multi-person system migrate, over years, into a single agent. The sections that follow trace that transfer one capacity at a time.

Identity assignment runs alongside the regulatory functions. The infant is given a name, a kinship position, a set of social categories, and a place in a family narrative, and it is held to expectations attached to these long before it can author any of them. The child learns who it is to others as the first available answer to who it is. This recovers, at the developmental base, the oldest layer of personhood in the anthropological record, the self as a node in kinship, ritual, and recognition rather than as a bounded interior given at birth (Mauss, 1985). The child is recognized into a position in a web of relations, and that recognized position is the scaffold on which a more individuated self is later built.

## 6. Joint attention and shared intentionality

Long before the first words, infants engage in proto-conversational exchanges of gaze, expression, and vocal turn-taking, the primary intersubjectivity that precedes and prepares joint attention (Trevarthen & Aitken, 2001). Around the end of the first year a structural change builds on it. Dyadic exchange, the back-and-forth of gaze and affect between infant and caregiver, becomes triadic. The infant follows a caregiver's gaze or point to an object, checks back to the caregiver's face, and coordinates attention to a shared target. The infant also begins to point declaratively, directing an adult's attention to something to share interest in it rather than to obtain it (Liszkowski et al., 2004). The dyad becomes a triangle of self, other, and world.

This triadic structure is the platform for word learning, teaching, and shared reference. To learn that a sound names a thing, the infant must know which thing the speaker means, and joint attention solves the reference problem by aligning what speaker and learner are attending to. Carpenter, Nagell, and Tomasello documented the emergence of joint-attention skills across the first eighteen months and their tight relation to early language, though joint attention is a platform rather than a single sufficient cause (Carpenter, Nagell & Tomasello, 1998). The social character of the input matters: infants who received live, interactive exposure to a foreign language showed phonetic learning that matched-content audio and video recordings failed to produce, which indicates that contingent social interaction rather than acoustic exposure alone drives early language learning (Kuhl, Tsao & Liu, 2003).

Tomasello's account places shared intentionality at the center of what makes the resulting cognition human. The child does not merely attend with others to objects. It enters "we" structures: shared goals, joint plans, divided roles, and a sense of how a task is done that is held in common and treated as binding. The progression runs from dyadic engagement through joint attention to collaborative activity in which each participant holds a role and tracks the other's (Tomasello et al., 2005; Tomasello, 2019). Ostensive communication trains the child to treat some signals as conveying generalizable knowledge rather than immediate commands: Csibra and Gergely's natural-pedagogy account argues that ostensive cues, direct gaze, infant-directed speech, and contingent responsivity, mark information as communicative and generic, which is what allows cultural knowledge to be transmitted rather than merely demonstrated (Csibra & Gergely, 2009). These perspective-taking abilities underlie the forms of social learning, imitative, instructed, and collaborative, that distinguish human cultural transmission (Tomasello, Kruger & Ratner, 1993).

The "we" layer is where individual agency becomes participation in shared structure. Cooperation, teaching, role division, and norm-governed joint action all presuppose that two agents can hold a goal in common and treat its standards as obligations on both. This is the developmental seed of the institutions that the later layers will install, because an institution is, at bottom, a shared intentional structure stabilized across many agents and inherited across generations.

## 7. Language as symbolic control architecture

Language is a control layer that reorganizes perception, memory, social life, and self-regulation. The child does not merely acquire labels for thoughts it already has. It acquires forms of thought that were not available before by entering a linguistic community, and the difference between the two readings is the difference between language as ornament and language as architecture.

The reorganizing functions can be itemized. Naming creates stable handles for entities, actions, states, and relations, so that the world becomes addressable. Categorization compresses perception into reusable concepts, so that "animal", "tool", "danger", and "promise" support inference beyond the instance. Self-reference, through the name and then the shifting pronouns, lets the organism become an object to itself: the child learns that "I" and "you" trade places depending on who speaks, which is a non-trivial discovery that the self is both a subject of experience and an item others refer to. Temporal vocabulary links past, present, and future, so that "I was", "I will", and "I promised" become statable. Causal and conditional morphology, "because", "if", "should have", enables explanation and counterfactual reasoning. Mental-state terms, "want", "think", "know", "remember", give the child a vocabulary for minds, its own and others'. And inner speech converts external dialogue into a medium of private thought.

Nelson's work on the mediated mind treats this transformation as the central event of early cognitive development. Language restructures memory, concept formation, narrative, and the understanding of other minds, because it supplies a semiotic medium in which experience can be represented, stored, recombined, and shared (Nelson, 1996). The claim is not that pre-linguistic cognition is empty. The claim is that a large class of distinctively human operations, autobiographical recall, explicit norm reasoning, decontextualized abstraction, public accountability, become available only once the child can operate on symbols.

The reach of the symbolic layer extends well past reference. Once thought can be borne in language it detaches from the perceptible situation: a child who has "yesterday", "if", and "because" can reason about what is absent, conditional, or counterfactual, and a community that has writing can externalize that reasoning into marks that outlast the reasoner and can be inspected, copied, and corrected by others. Inner speech is the hinge. The verbal self-regulation that begins as audible private speech becomes a silent medium in which the agent poses questions to itself, holds a premise while testing a consequence, and rehearses an argument before committing to it. Scientific and mathematical reasoning sit at the far end of this trajectory. Explicit definitions, public hypotheses, and the obligation to be refutable are operations on symbols that the bare perceptual system cannot perform, and they are available only to an agent that has internalized a language and the collective practices built on top of it.

The boundary case that fixes the claim is late first-language acquisition. Mayberry and Lock studied deaf and hearing adults who differed in whether they had accessible language in early childhood, and found that those with early first-language exposure, signed or spoken, later achieved high competence in a second language, while those who lacked accessible language in the first years performed poorly even after extended later exposure (Mayberry & Lock, 2003). The critical

variable is accessibility rather than modality: sign language acquired early supports later language as well as speech does, and speech that the child cannot hear supports nothing. The finding establishes that early accessible language is developmental infrastructure rather than a late addition, because its absence during the sensitive window cannot be fully repaired afterward. This is why severe language deprivation is so damaging. Language is one of the central media through which human cognitive control becomes recursive, and a window missed is a layer that the later layers cannot fully assume.

## 8. Norm internalization and moral agency

A person in the cognitive sense is an agent whose actions are answerable. Preference and aversion are not yet conduct. Conduct is behavior interpreted through shared standards, and the capacity to treat one's own behavior as answerable to such standards is acquired, like the others, from outside in.

The acquisition begins with correction. Caregivers and peers prohibit, permit, praise, and blame, and the child first encounters norms as the external consequences of acts. The developmental step that matters is the move from "this brings punishment" to "this is forbidden", because the second generalizes beyond the immediate enforcer to a rule that holds in a context. Children make this move early and robustly. By the third year they not only follow norms but enforce them on third parties, spontaneously protesting and correcting a puppet's violations of a game's rules even when they are not themselves affected, and treating the norm as something that binds anyone in the relevant situation (Rakoczy, Warneken & Tomasello, 2008). Enforcement of a rule one is not party to is strong evidence that the rule has been internalized as a general standard rather than learned as a personal contingency. The same norm sensitivity shows in how children copy: they faithfully reproduce even the causally unnecessary steps of a demonstrated procedure, an overimitation that treats the modeled sequence as the right way to do the thing, and a high-fidelity copying that separates human children from other apes (Lyons, Young & Keil, 2007; Horner & Whiten, 2005).

Thompson's developmental account links this to the broader growth of conscience, embedding norm sensitivity in the same matrix as social understanding, relationships, and the self (Thompson, 2006). The self-conscious emotions of accountability, guilt, shame, and pride, develop alongside the rules and presuppose both a represented self and a standard to measure it against; their onset is tied to the objective self that emerges with mirror self-recognition in the second year (Lewis et al., 1989). They are what convert a violated standard into an experienced obligation. A child who feels the pull to apologize after a rupture has internalized not only the rule but the stance of being answerable for breaking it.

The layering of normative scope continues well past early childhood. Family norms give way to peer norms, peer norms to institutional norms, and institutional norms to the explicit, codified, jurisdiction-specific norms of law and profession. Each widening reuses the same machinery: a standard is encountered as external enforcement, abstracted into a general rule, bound to an emotion of accountability, and folded into the self-model as a constraint the agent imposes on itself. Moral personhood is the cumulative result, an agent whose action selection is shaped by standards

it treats as binding independent of who is watching.

## 9. Autobiographical memory and narrative selfhood

The narrative self is co-authored before it is self-authored. A young child has event memory and procedural memory, but it does not yet have an organized, owned, temporally ordered story of its own life. That story is assembled in conversation with adults who remember on the child's behalf and, in remembering aloud, teach the child the form a life takes.

The mechanism is elaborative reminiscing. Caregivers revisit shared events with the child in a particular grammar: who was there, what happened, what the child felt, why it mattered, what kind of person the child was in the event. Nelson and Fivush's social-cultural theory holds that autobiographical memory emerges across the preschool years from the convergence of basic memory capacity, language, adult memory talk, temporal understanding, and a developing sense of self and other (Nelson & Fivush, 2004). Caregivers who reminisce in an elaborative style, asking open questions and richly reconstructing past episodes, have children who develop more detailed and coherent autobiographical narratives, an association that the longitudinal evidence suggests runs at least partly from the conversational style to the child's later memory rather than the reverse (Reese, Haden & Fivush, 1993). The child learns to remember in the first person by being walked through first-person remembering until it can do the walking alone.

Childhood amnesia is the clue that fixes the construction. Adults retain almost no episodic memories from before roughly the third or fourth year, and the offset of the amnesia tracks the development of language and narrative skill rather than the development of memory storage as such. The pattern is what the co-authoring account predicts: episodic events occur and are stored before the child has the narrative apparatus to organize them into a retrievable autobiography, and what is lost is not the events but the indexing system that would have made them tellable.

Narrative selfhood is the result of this process running for years. The self becomes a temporally extended protagonist, an agent who did things, to whom things happened, who has a trajectory and a character. The family supplies the first version of the trajectory and the first reading of the character, and the child inherits both before revising them. The narratable individual, the named self with a tellable life course, depends on recognition by others throughout: a life becomes a story when someone has been telling it.

## 10. Executive function, private speech, and inner speech

Adult agency requires capacities that the young child does not have and the caregiver initially supplies: holding a goal in mind, inhibiting a prepotent response, sequencing steps, correcting errors, switching between tasks. These executive functions develop steeply across the preschool and school years, and their development is partly the internalization of an external supervisor.

The route is the one Vygotsky described: higher mental functions appear first between people and only later within the child, as interpsychological regulation becomes intrapsychological (Vygotsky, 1978). The adult regulates the child's action from outside through speech, the kind of

graded support that lets a child accomplish what it cannot yet manage alone, termed scaffolding by Wood, Bruner, and Ross (1976): "wait", "first this, then that", "don't touch", "try again", "use your words". The child then reproduces this speech aloud, talking itself through tasks in audible self-directed private speech, and finally compresses the private speech into silent inner speech that guides action from within, the trajectory from egocentric to inner speech that Vygotsky (1987) made central to verbal thought. The trajectory is well attested: private speech increases when tasks are difficult, predicts better performance, and is gradually internalized into covert verbal self-regulation, so that social command becomes self-command (Winsler, Fernyhough & Montero, 2009). The supervisor that once stood outside the child comes to operate inside it, in the same verbal medium it used when it was external.

The executive layer is the clearest case of the architecture's central pattern, and it shows why the pattern is more than metaphor. There is a measurable transfer of a regulatory function from a two-person system to a one-person system, mediated by a public representational medium that becomes private. Working memory, inhibition, and cognitive flexibility do have a maturational component grounded in prefrontal development, and that maturation supplies only the substrate. In humans the content of self-regulation, the specific instructions the child gives itself, the order in which it does things, the standards against which it checks its work, is drawn from the social regulation it received, then run on maturing neural hardware. School formalizes and intensifies this transfer, imposing schedules, routines, and explicit demands for sustained attention and delayed action that train executive control well beyond what unstructured life requires.

## 11. Cumulative culture and institutional personhood

Most developmental accounts stop in childhood. The bootstrapping architecture does not, because the higher layers of personhood are installed across adolescence and adulthood and depend on resources that no individual could generate alone. The adult is a nervous system coupled to cumulative culture, and the coupling is what produces the capacities that look most like distinctively human personhood.

Cumulative culture is the inheritance that makes this possible. Human populations preserve and modify innovations across generations through high-fidelity social transmission, so that each learner starts from an accumulated stock of tools, symbols, procedures, and institutions that took many lifetimes to build. This ratchet, in which faithful copying prevents loss and frees variation to accumulate, runs at a scale no other species matches (Tomasello, 1999; Tennie, Call & Tomasello, 2009), and it is the substrate of the dual inheritance, genetic and cultural, that shaped human cognition (Boyd & Richerson, 1985; Henrich, 2016). The child does not reinvent number, writing, law, or scientific method. It inherits them, compressed into transmissible form, and internalizes them through the same participation-then-internalization loop that governed the earlier layers, now operating on cultural artifacts rather than on caregiver behavior.

The institutional layers can be itemized by what each installs. Literacy externalizes memory into durable marks, so that a person can think with more than working memory holds and can stand in relation to texts produced by the dead. Numeracy supplies a compression system for quantity

and relation that perception does not provide. Schooling shapes attention, imposing sustained focus, sequenced curricula, and the habit of treating decontextualized symbols as worth attending to. Law supplies a public grammar of personhood, defining who may own, contract, testify, and be held answerable, in the direct institutional descent from the Roman *persona*. Money and identity documents make the person legible to economic and bureaucratic systems, so that the self acquires a representation that institutions can read and act on. Science installs the norms of a collective self-correcting practice, explicit hypotheses, public evidence, and the obligation to be refutable. Professions package role-specific competence, obligation, and identity into inhabitable forms. Each of these is a cultural structure first run around the developing agent and then partly internalized as a capacity and a self-description.

This is how the same initial architecture diverges into a deprived survivor or a research scientist. The difference is principally the quantity and quality of external cognitive structure that the agent has had the opportunity to internalize, far more than brain size or general intelligence. Education exerts a measurable cognitive effect of the kind the architecture predicts: a meta-analysis of 142 effect sizes from 42 datasets and over 600,000 participants, drawing on three quasi-experimental designs, estimated that each additional year of schooling raises measured intelligence by roughly one to five IQ points, an effect that showed little sign of fading with age (Ritchie & Tucker-Drob, 2018). The adult scientist is the infant plus an internalized civilization, and the institutions are constitutive of the highest layers of the mind they help to build rather than decoration on a mind already finished without them.

The history of the concept of the person supports reading these layers as a genuine stratification. Mauss's classic argument treats "the person" as a category assembled over the long history of societies, moving from the relational and ritual personage through the Roman juridical *persona* to the interior moral self and finally the rational, rights-bearing individual (Mauss, 1985). The developmental architecture recapitulates that stratification under compression. The relational layer is installed in infancy through co-regulation and social address; the narratable individual is assembled in the autobiographical self co-authored across early childhood; the juridical and rational-autonomous layers arrive with the institutional and reflective capacities installed by schooling, law, and disciplined inquiry across adolescence and adulthood. History assembles the layers across many generations. Development runs the same layering inside a single life, which is why disassembly at either scale, historical or biographical, follows the same logic: withdraw the institutional supports and the layer they sustain thins, while the formal category persists.

## 12. Deprivation as a boundary condition

Deprivation is the closest thing developmental science has to an experiment on the architecture, and it is an ethically and methodologically compromised one. The cases never isolate a single missing input, the documentation is often poor, and the children involved are frequently subject to multiple overlapping harms. The evidence nonetheless converges on a single structural conclusion, which is what gives it weight despite the noise.

Three lines of evidence converge. Institutional deprivation supplies the first. The Bucharest Early

Intervention Project, a randomized controlled trial of foster-care placement for institutionally reared children, found that early institutional care, characterized by low responsive caregiving and impoverished stimulation, depressed cognitive development, and that placement into responsive foster care, particularly before roughly two years of age, substantially improved cognitive outcomes (Nelson et al., 2007). Follow-ups in the same cohort traced the deficits into memory and executive function (Bos et al., 2009) and into brain electrical activity, which partially normalized with earlier placement (Vanderwert et al., 2010). The independent English and Romanian Adoptees study found that deprivation lasting beyond the first six months left neurodevelopmental and mental-health effects detectable into young adulthood (Sonuga-Barke et al., 2017). The dose-and-timing structure of these results is the signature of an experience-expectant system (Greenough, Black & Wallace, 1987): the deficits track the absence of expected caregiving input and are partly remediable if the input arrives within a window. Late first-language acquisition supplies the second line, established above: accessible language missing in the early years cannot be fully compensated later, which marks language as infrastructure rather than enrichment (Mayberry & Lock, 2003). Language deprivation in deaf children who lack early access to sign supplies a clinical version of the same finding, with patterns of dysfluent language, reduced general knowledge, and disrupted reasoning described under the heading of language-deprivation effects (Hall, 2017).

Extreme isolation cases, the so-called feral and confined children, are the third line and the weakest. They are rare, almost always confounded by abuse, malnutrition, and possible pre-existing disability, and frequently sensationalized in ways that make clean inference impossible. The best-documented modern case, Genie, a girl discovered at thirteen after years of confinement and isolation, acquired substantial vocabulary but never normal grammar, which is consistent with a sensitive-period limit on first-language acquisition, though her history of abuse and possible prior disability confound any single interpretation (Curtiss, 1977). Reviews of the broader material warn that such cases are too poorly documented to settle questions about human nature on their own (Candland, 1993). They are best treated as suggestive corroboration of conclusions established by the better-controlled evidence; on their own they settle nothing.

The correct reading of all three lines is a single careful sentence. Deprivation does not reveal what a human is in the absence of culture. It reveals that the human organism expects a social, linguistic, and cultural niche for normal personhood development, and that the architecture fails in characteristic ways when the expected niche is withheld. What remains under deprivation is the proto-agent and its local adaptations: affect, sensorimotor skill, nonverbal memory, attachment-seeking in distorted forms, practical problem-solving. What fails is the upper stack: language-mediated cognition, narrative selfhood, explicit norm reasoning, and institutional competence. The pattern of sparing and loss is the architecture's strongest confirmation, because it is what a layered, experience-expectant bootstrapping system predicts and what neither pure nativism nor pure environmentalism predicts cleanly.

## 13. A bootstrapping architecture

The mechanisms above compose into one architecture. The following table maps each mechanism to the external scaffold that supplies it, the capacity the child internalizes, and the contribution to personhood, which makes the recursive structure legible at a glance.

| Layer | External scaffold | Internalized capacity | Contribution to personhood |
|---|---|---|---|
| Affective regulation | Feeding, holding, soothing, repair | Self-soothing, affect tolerance | Stability of agency |
| Attachment | Secure base, safe haven | Trust, exploration, relational priors | A social world that can be navigated |
| Attention regulation | Pointing, showing, gaze-following | Joint attention, relevance tracking | A world held in common |
| Emotion concepts | Marked affect mirroring, labeling | Emotion differentiation | A mental-state reading of the self |
| Language | Naming, requests, stories, questions | Categories, self-reference, inner speech | A symbolically structured self |
| Shared intentionality | Joint action, divided roles | "We" structures, cooperation | Participation in shared agency |
| Norms | Correction, praise, blame, permission | Rule-following, accountability | Moral personhood |
| Autobiographical memory | Elaborative reminiscing | Narrative continuity | A self extended across time |
| Executive function | Adult planning, reminders, private speech | Inhibition, planning, self-command | Deliberate agency |
| Cumulative culture | Literacy, numeracy, law, science, roles | Cultural tools and institutional competence | Institution-compatible adulthood |

The architecture is a loop, and its inputs, scaffolds, internalization processes, and outputs can be stated as a conceptual schema. The pseudocode below describes the developmental process at a high level; it is neither an executable model nor a claim about neural implementation. Its purpose is to make the recursion and the attractor structure explicit.

```
Initialize newborn:
    body          = embodied_homeostatic_system()
    affect        = active(distress, comfort, interest, arousal)
    world_model   = shallow, rapidly plastic
    caregiver     = uncalibrated prior over availability
```

```
    social_bias   = oriented to faces, voices, biological motion
    language      = no lexicon, strong statistical-learning capacity
    self_model    = minimal embodied self, no narrative identity
    norms         = minimal
    autobiography = minimal
    executive     = externally scaffolded
    culture       = none internalized

For each developmental interaction:

    perceive(body, world, caregiver, symbols)
    predict(causes, caregiver_response, action_outcomes)
    act(cry, orient, reach, gaze, vocalize, imitate, ask, protest)

    if distress > self_regulation_capacity:
        signal(caregiver)
        receive(comfort, feeding, voice, repair)
        update(caregiver_prior, affect_regulation)

    if caregiver marks/labels state:
        bind(bodily_affect_pattern -> emotion_concept)

    if joint_attention(self, caregiver, object):
        align(attention); create triadic reference

    if language_input under social contingency:
        extract(statistical_structure)
        map(words -> objects, actions, states, relations)
        learn(names, pronouns, tense, causality, mental-state terms)

    if norm_feedback:
        abstract(enforcement -> general rule in context)
        bind(rule -> accountability emotion)

    if autobiographical_talk:
        encode(event + emotion + agent + cause as self-in-time)

    if private_speech_practice:
        internalize(external_instruction -> self_instruction -> inner_speech)

    if education_or_institution:
```

```
        internalize(cultural_tools, roles, abstract_symbols)

    self_model = integrate(body, name, memory, language,
                           norms, roles, projects)

    # recursion: this round's new capacity changes next round's data
    caregiver_and_peers.adjust_to(self_model)
```

The attractor logic follows from the loop. Where the environment supplies responsive caregivers, accessible language, stable attachment figures, joint attention, play, norms with repair, education, and cultural tools, the system converges on increasing self-regulation, language-mediated cognition, narrative selfhood, norm internalization, institutional compatibility, and reflective agency. Where the environment lacks responsive care, accessible language, stable social feedback, emotional labeling, norm scaffolding, and education, the system preserves basic organism and proto-agent functions and local survival adaptations, while the upper layers, language-mediated cognition, narrative selfhood, executive and institutional capacity, fail to develop normally. The two trajectories are one algorithm run on two niches, which is why the boundary cases of section 12 are evidence for the architecture rather than exceptions to it.

## 14. What the architecture is consistent with, what it predicts, and what it declines to claim

Honesty about an integrative account requires separating two things that are easy to conflate. Most of what an assembled synthesis can offer is consistency: it must accommodate the findings it was built from, and pointing to that accommodation is not prediction but retrodiction. A smaller class of claims is genuinely predictive, in that the architecture commits to a relationship that its component literatures do not each separately assert and that a determined critic could find false. This section keeps the two apart, because the paper's value as an integration and its much weaker standing as a falsifiable theory are different questions and deserve different answers.

The following are consistency claims rather than predictions. Each restates, in the architecture's vocabulary, a relationship that the studies marshalled earlier already report, and listing them confirms only that the synthesis is faithful to its sources. That elaborative parental reminiscing accompanies more coherent autobiographical narrative is an established finding of the reminiscing literature rather than a forecast the architecture risks (Reese, Haden & Fivush, 1993; Nelson & Fivush, 2004). That the private-speech-to-inner-speech transition tracks executive-function gains is the established reading of the private-speech literature (Winsler, Fernyhough & Montero, 2009). That severe early language deprivation impairs the symbolic upper stack while sparing more nonverbal capacities is the pattern already drawn from the deprivation and late-acquisition evidence (Mayberry & Lock, 2003; Hall, 2017). That more schooling accompanies measurably more abstract cognition is the finding of the education meta-analysis already cited (Ritchie & Tucker-Drob, 2018). These relationships are why the architecture was assembled the way it was; they cannot then be offered as its successful predictions without circularity.

A narrower set of claims is closer to genuinely predictive, because it concerns relationships that follow from the integration but are not the direct content of any single assembled study. The architecture entails specific mediation structure, and mediation is harder to retrodict than correlation. If self-repair is internalized dyadic repair, then the link between caregiver repair after rupture and the child's later capacity for apology and error correction should be mediated by the child's acquisition of repair scripts in private and inner speech, and should not reduce to general attachment security. If triadic reference is the platform for symbolic self-development specifically, then early joint-attention quality should predict later pronoun mastery and narrative ability over and above its prediction of vocabulary size, with which it is already known to correlate. The architecture's distinctive layering claim, that each layer rewrites the learning conditions for the next, predicts ordering and gating effects rather than mere co-occurrence: emotion-labeling density should affect later emotion regulation through, and not independently of, the child's emotion-concept vocabulary, and a layer should be hard to install when the layer below it is absent. These are the claims on which the architecture could actually be tested, and the paper offers them as the part of the synthesis that earns the word prediction. They are directional and stated as relationships under reasonable controls rather than as guarantees, and several remain untested in the specific mediated form the architecture requires.

The model's value depends equally on what it declines to claim. It does not claim that infants lack moral worth; the developmental and moral senses of personhood are held separate by construction. It does not claim that culture writes on a blank slate; the newborn is an active, biased, prediction-driven learner, and temperament, sensitive periods, genetics, and embodiment constrain every trajectory. It does not claim that language is necessary for consciousness or experience; the claim is the narrower one that language is necessary for the higher symbolic, narrative, and accountable layers. It does not claim that caregivers install a self into passive hardware; the child actively recruits, predicts, imitates, resists, and explores, and its signals shape the very environment that scaffolds it. It does not treat feral-child cases as probative; they corroborate, weakly, conclusions established elsewhere. And it does not read the "external runtime" and "symbolic operating system" descriptions as literal mechanisms; they are explanatory heuristics for a transfer of regulatory and representational functions from a multi-person system to a single agent, and that transfer is the claim the metaphor stands in for.

A final limit concerns the status of the architecture itself, and it should be conceded rather than managed. The account is an integrative synthesis. Its contribution is the demonstration that mechanisms studied in separate literatures can be staged as a single recursive bootstrapping stack running from the neonate to the institution-compatible adult, and almost everything it accommodates it accommodates because it was assembled to. This carries a real falsifiability cost. A framework flexible enough to absorb co-regulation, language, norms, narrative, executive control, and cumulative culture is, by that flexibility, hard to refute as a whole, and much of what reads as confirmation is consistency with studies already in hand rather than risky prediction borne out. The architecture is differentiated from the interactionist consensus it shares by its specific staging and its reach into the institutional layers, not by a novel mechanism, and a critic is entitled to ask whether the staging carves development at real joints or merely imposes a tidy order on a continuous process.

The honest defense is modest: the synthesis is worth keeping if its layered, experience-expectant structure organizes the evidence more economically than the alternatives and if the mediation and ordering claims of the preceding paragraph survive being tested in the specific form the architecture requires. Where those claims fail, the architecture should be revised or abandoned, and the integration's convenience is not evidence in its favor.

## 15. The culture-completable organism

The human infant is born as an embodied proto-agent: a homeostatic, affective, sensorimotor, socially oriented, plastic prediction-and-control system with a structural expectation of external regulation. Adult cognitive personhood is bootstrapped. The interactionist premise this shares with neuroconstructivism and developmental systems theory, that development emerges from organism-environment coupling rather than from a maturational timetable or inscription onto a passive surface, is not the contribution; the contribution is the specific staging of that coupling and its extension into the institutional layers. The infant enters caregiver-mediated loops of co-regulation, social address, joint attention, language, norm enforcement, autobiographical narration, and executive scaffolding, and across years it internalizes the regulatory and symbolic structures first run around it by others, until control that began as a property of a two-person system becomes a property of one agent. Language reorganizes the architecture into a recursive one, memory binds the self across time, norms make action answerable, and cumulative culture installs the institutional and reflective layers that no individual could generate alone.

Human personhood, in the developmental-cognitive sense, is co-regulated, language-mediated, norm-governed, narratively extended, and culturally completed agency. The newborn is a culture-completable organism, and the completion is a real developmental process with identifiable mechanisms, identifiable failure modes under deprivation, and testable consequences. The layered structure of this developmental account mirrors the layered structure that the long history of the person concept already records, and it bears the same vulnerability: withdraw the supports that sustain a layer and the layer thins, whatever the formal category continues to assert.

## References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of Attachment: A Psychological Study of the Strange Situation*. Lawrence Erlbaum.

Atzil, S., Gao, W., Fradkin, I., & Barrett, L. F. (2018). Growing a social brain. *Nature Human Behaviour*, 2(9), 624-636.

Bos, K. J., Fox, N., Zeanah, C. H., & Nelson, C. A. (2009). Effects of early psychosocial deprivation on the development of memory and executive function. *Frontiers in Behavioral Neuroscience*, 3, 16.

Bowlby, J. (1969). *Attachment and Loss, Vol. 1: Attachment*. Basic Books.

Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press.

Candland, D. K. (1993). *Feral Children and Clever Animals: Reflections on Human Nature*. Oxford University Press.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4, Serial No. 255).

Cittern, D., Nolte, T., Friston, K., & Edalat, A. (2018). Intrinsic and extrinsic motivators of attachment under active inference. *PLOS ONE*, 13(4), e0193955.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181-204.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, 61(5), 1584-1595.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), 148-153.

Curtiss, S. (1977). *Genie: A Psycholinguistic Study of a Modern-Day "Wild Child."* Academic Press.

Davis, J., Redshaw, J., Suddendorf, T., Nielsen, M., Kennedy-Costantini, S., Oostenbroek, J., & Slaughter, V. (2021). Does neonatal imitation exist? Insights from a meta-analysis of 336 effect sizes. *Perspectives on Psychological Science*, 16(6), 1373-1394.

De Wolff, M. S., & van IJzendoorn, M. H. (1997). Sensitivity and attachment: A meta-analysis on parental antecedents of infant attachment. *Child Development*, 68(4), 571-591.

Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2002). *Affect Regulation, Mentalization, and the Development of the Self*. Other Press.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127-138.

Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, 4(1), 14-21.

Gergely, G., & Watson, J. S. (1996). The social biofeedback theory of parental affect-mirroring: The development of emotional self-awareness and self-control in infancy. *International Journal of Psycho-Analysis*, 77(6), 1181-1212.

Greenough, W. T., Black, J. E., & Wallace, C. S. (1987). Experience and brain development. *Child Development*, 58(3), 539-559.

Griffiths, P. E., & Gray, R. D. (1994). Developmental systems and evolutionary explanation. *The Journal of Philosophy*, 91(6), 277-304.

Hall, W. C. (2017). What you don't know can hurt you: The risk of language deprivation by impairing sign language development in deaf children. *Maternal and Child Health Journal*, 21(5), 961-965.

Henrich, J. (2016). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition*, 8(3), 164-181.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40(1-2), 1-19.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. MIT Press.

Köster, M., Kayhan, E., Langeloh, M., & Hoehl, S. (2020). Making sense of the world: Infant learning from a predictive processing perspective. *Perspectives on Psychological Science*, 15(3), 562-571.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, 100(15), 9096-9101.

Lewis, M., & Brooks-Gunn, J. (1979). *Social Cognition and the Acquisition of Self*. Plenum Press.

Lewis, M., Sullivan, M. W., Stanger, C., & Weiss, M. (1989). Self development and self-conscious emotions. *Child Development*, 60(1), 146-156.

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science*, 7(3), 297-307.

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences*, 104(50), 19751-19756.

Mareschal, D., Johnson, M. H., Sirois, S., Spratling, M. W., Thomas, M. S. C., & Westermann, G. (2007). *Neuroconstructivism: How the Brain Constructs Cognition* (Vol. 1). Oxford University Press.

Mauss, M. (1985). A category of the human mind: The notion of person; the notion of self (W. D. Halls, Trans.). In M. Carrithers, S. Collins, & S. Lukes (Eds.), *The Category of the Person: Anthropology, Philosophy, History* (pp. 1-25). Cambridge University Press. (Original work published 1938)

Mayberry, R. I., & Lock, E. (2003). Age constraints on first versus second language acquisition: Evidence for linguistic plasticity and epigenesis. *Brain and Language*, 87(3), 369-384.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2), 143-178.

Meltzoff, A. N., Murray, L., Simpson, E., Heimann, M., Nagy, E., Nadel, J., Pedersen, E. J., Brooks, R., Messinger, D. S., De Pascalis, L., Subiaul, F., Paukner, A., & Ferrari, P. F. (2018). Re-examination of Oostenbroek et al. (2016): Evidence for neonatal imitation of tongue protrusion. *Developmental Science*, 21(4), e12609.

Nelson, C. A., Zeanah, C. H., Fox, N. A., Marshall, P. J., Smyke, A. T., & Guthrie, D. (2007). Cognitive recovery in socially deprived young children: The Bucharest Early Intervention Project. *Sci-*

*ence*, 318(5858), 1937-1940.

Nelson, K. (1996). *Language in Cognitive Development: The Emergence of the Mediated Mind*. Cambridge University Press.

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review*, 111(2), 486-511.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255-258.

Oyama, S. (1985). *The Ontogeny of Information: Developmental Systems and Evolution*. Cambridge University Press.

Oostenbroek, J., Suddendorf, T., Nielsen, M., Redshaw, J., Kennedy-Costantini, S., Davis, J., Clark, S., & Slaughter, V. (2016). Comprehensive longitudinal study challenges the existence of neonatal imitation in humans. *Current Biology*, 26(10), 1334-1338.

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology*, 44(3), 875-881.

Reese, E., Haden, C. A., & Fivush, R. (1993). Mother-child conversations about the past: Relationships of style and memory over time. *Cognitive Development*, 8(4), 403-430.

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science*, 29(8), 1358-1369.

Rochat, P. (2003). Five levels of self-awareness as they unfold early in life. *Consciousness and Cognition*, 12(4), 717-731.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928.

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences*, 105(2), 809-813.

Sonuga-Barke, E. J. S., Kennedy, M., Kumsta, R., Knights, N., Golm, D., Rutter, M., Maughan, B., Schlotz, W., & Kreppner, J. (2017). Child-to-adult neurodevelopmental and mental health trajectories after early life deprivation: The young adult follow-up of the longitudinal English and Romanian Adoptees study. *The Lancet*, 389(10078), 1539-1548.

Taipale, J. (2016). Self-regulation and beyond: Affect regulation and the infant-caregiver dyad. *Frontiers in Psychology*, 7, 889.

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B*, 364(1528), 2405-2415.

Thompson, R. A. (2006). The development of the person: Social understanding, relationships, conscience, self. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Psychology* (6th ed., Vol. 3, pp. 24-98). Wiley.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press.

Tomasello, M. (2019). *Becoming Human: A Theory of Ontogeny*. Harvard University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), 675-735.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16(3), 495-552.

Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry*, 42(1), 3-48.

Vanderwert, R. E., Marshall, P. J., Nelson, C. A., Zeanah, C. H., & Fox, N. A. (2010). Timing of intervention affects brain electrical activity in children exposed to severe psychosocial neglect. *PLoS ONE*, 5(7), e11415.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, 10(2), 159-164.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.

Vygotsky, L. S. (1987). Thinking and speech (N. Minick, Trans.). In R. W. Rieber & A. S. Carton (Eds.), *The Collected Works of L. S. Vygotsky, Vol. 1: Problems of General Psychology* (pp. 39-285). Plenum Press. (Original work published 1934)

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3), 655-684.

Winsler, A., Fernyhough, C., & Montero, I. (Eds.). (2009). *Private Speech, Executive Functioning, and the Development of Verbal Self-Regulation*. Cambridge University Press.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89-100.