

Assessment-First in Early Childhood Science Education: An Exploratory Practice of an Integrated “Intelligent Assessment-Intelligent Teaching-Intelligent Learning” Closed-Loop Model in a Kindergarten Junior Class

Manting Ke¹

¹ Tianlu Kindergarten, Huangpu District, Guangzhou, China

Correspondence: Manting Ke, Tianlu Kindergarten, Huangpu District, Guangzhou, China.

doi:10.63593/JARE.2026.09.03

Abstract

China’s current reform of educational evaluation has drawn renewed attention to a persistent problem in early childhood science education: teaching, learning, and assessment have long functioned as separate processes. The signs are ordinary enough: activities planned from experience, inquiry that stays near the surface, judgment reserved for the end. Our study asks whether the order can be reversed—whether assessment, moved to the front, can pull teaching and learning together. The model we propose rests on two bodies of research, one on children’s preconceptions and one on deep learning. Intelligent assessment, intelligent teaching, and intelligent learning are joined in a data-driven loop that spirals upward; diagnosis opens the process and keeps informing it. We implemented the model in four steps over two months with 22 children in a junior class (ages 3–4), under the science theme “The Big Tree and Its Friends.” Digital tools supported the elicitation interviews, learning analytics, and process documentation throughout. The pre- and post-intervention evidence is encouraging. Recognition of concealed tree structures rose from 22.7% to 88.0%, and the proportion of children able to explain, in scientific terms, what plants need to grow rose from 9.1% to 90.9%, with visible gains in the quality of inquiry as well. These results suggest that moving assessment to the front of instruction can turn evaluation from a terminal verdict into navigational guidance at the very start of learning, while also helping teachers grow into data-informed, research-oriented practitioners and drawing families into closer cooperation with the kindergarten. For grassroots settings, the study offers a workable route for putting evaluation reform into practice.

Keywords: educational assessment, assessment-first, early childhood science education, deep learning, data-driven instruction, formative assessment

1. Introduction

Educational evaluation works as a central

steering mechanism in the construction of a high-quality education system. The Overall Plan for

Deepening Educational Evaluation Reform in the New Era (CPC Central Committee and State Council, 2020) laid out a multi-dimensional evaluation framework and carried China's evaluation reform into a stage where dismantling outdated practices and building new ones go hand in hand, with construction given the priority. How a macro-level policy of this kind can be translated into everyday classroom practice—and what evaluation should, in the end, contribute to educational modernization—is where the present study begins.

Science education in the early years is foundational to the cultivation of innovative talent, and the kindergarten years are precisely the window in which dispositions toward scientific inquiry first take shape (National Research Council, 2007). Evaluation practice in preschools, however, still leans heavily on summative, outcome-oriented judgment. This kind of practice overlooks the dynamic character of young children's development and the sheer size of individual differences, and it fits awkwardly with a development-centered orientation. Curriculum evaluation is weakened further by vaguely worded objectives, fragmented content, and thin practical enactment, while evaluation standards and instrument development lag behind. On the front line of practice, these systemic tensions condense into three problems that recur again and again.

The first is a misalignment between instructional design and children's actual learning states. Objectives are often set from adult logic and generic experience, which rarely matches the imaginative, pre-conceptual world of a young child; teaching then proceeds in a one-directional manner, with limited effect. A second difficulty lies in the inquiry itself. Young children observe with the naked eye, and what the naked eye cannot reach—the mechanism behind a phenomenon—usually stays hidden from them. So, their inquiry often stops at describing what lies on the surface, and these fleeting impressions rarely settle into the stable, core experiences we hope children will carry with them. The third difficulty is about timing. Assessment usually arrives at the very end of a lesson, which means that while children are actually inquiring, no feedback reaches them. A judgment made only at the end cannot register the confusion, or the small breakthroughs, that occur along the way. It becomes little more than a stamp on the final product, and teaching, learning, and assessment

drift further apart.

These problems pushed us back to a question that is easier to ask than to answer: what if assessment did not wait at the finish line, but stood at the starting point instead? For the frame we drew on deep learning theory, though what we had watched in real kindergarten classrooms shaped the model at least as much. In that model, assessment moves to the front of instruction and stays active there, running as a data-supported, spiraling loop through which intelligent assessment, intelligent teaching, and intelligent learning continually inform one another. If it works, the three strands are reconnected, and frontline teachers gain a model they can make their own.

2. Theoretical Foundations and Framework Construction

2.1 Policy and Theoretical Basis

Two policy documents frame this study: the Overall Plan for Deepening Educational Evaluation Reform in the New Era (CPC Central Committee and State Council, 2020) and the Guidelines for Learning and Development of Children Aged 3–6 (Ministry of Education of the People's Republic of China, 2012). On the theoretical side, we draw on two bodies of work—studies of preconceptions and studies of deep learning. As that literature describes it, deep learning does not proceed in a straight line. It starts from an effort to diagnose and restructure learners' existing ideas, develops through embodied inquiry situated in authentic contexts, and climbs in a spiral rather than along a linear path (Marton & Säljö, 1976; Biggs & Tang, 2011). This picture of learning is precisely what justifies running an integrated assessment loop through the entire instructional cycle.

2.2 Preconceptions: The Logical Starting Point and Diagnostic Basis of Learning

Before formal science instruction ever begins, learners have already built intuitive pictures of the natural world out of everyday experience, spontaneously and without guidance (Vosniadou & Brewer, 1992). A child's mind is no blank slate. These early understandings form the child's own framework for making sense of the world—a framework that resists change, depends heavily on context, and yet, for the child, carries real explanatory power (Carey, 1985).

A preschooler's conviction that a big tree needs to "drink milk" is a typical anthropomorphic

preconception of this kind. Traditional instruction tends either to ignore such ideas or to try to overwrite them directly with scientific concepts, yet decades of conceptual-change research suggest that this accomplishes little: old and new conceptions usually coexist and compete rather than one substituting for the other (Posner, Strike, Hewson, & Gertzog, 1982).

Recognizing and diagnosing children's preconceptions is therefore the logical starting point of science teaching—and the fundamental reason this study places assessment before instruction. Ausubel (1968) put it plainly: the most important single factor influencing learning is what the learner already knows. If the imaginative “theories” already at work in children's minds cannot be accurately identified, any instructional design is shooting in the dark. Front-loaded assessment makes these tacit, highly individualized preconceptions visible, and in doing so tells the teacher exactly where to intervene.

2.3 Deep Learning Theory: The Process and Destination of Learning

Deep learning is not the shallow memorization of knowledge. It denotes the learner's critical engagement with new ideas, the weaving of those ideas into existing cognitive structures, and the transfer of knowledge to novel situations (Marton & Saljo, 1976). In essence, deep learning happens while preconceptions are being challenged, revised, and rebuilt. On the pathway of this process the literature broadly converges: activating prior conceptions, provoking cognitive conflict, constructing scientific concepts, and applying them through transfer.

The critical lever in this cycle is the provocation of cognitive conflict. When children discover through embodied inquiry—handling materials, or watching phenomena that contradict their expectations—that what they previously believed can no longer account for the situation in front of them, genuine conceptual construction becomes possible (Posner et al., 1982). What this means for teaching is fairly direct. The core task is no longer one-way transmission; it is to design authentic situations and inquiry tasks that can provoke productive cognitive conflict, so that children do the meaning-making themselves as they travel from naive ideas toward scientific concepts. Seen this way, assessment also falls into line with learning—exactly as the formative assessment literature has long argued. Evidence

of prior understanding is gathered not to certify children, but to steer the teaching that follows (Black & Wiliam, 1998; Wiliam, 2011).

2.4 Framework Construction: The Three-Loop, Four-Step Dynamic Assessment Model

On these foundations we built what we call a “three-loop, four-step” dynamic assessment model aimed at promoting deep learning (see Figure 1). In this model, assessment serves as the engine. It drives teaching and learning forward, and the three together form a data-supported, spiraling closed system—one that puts into practice a simple logic: diagnose preconceptions first, design for cognitive conflict, then support deep construction.

The model runs in four steps.

The first step happens before any teaching at all. With the help of digital tools, we collect children's original understandings of the inquiry topic and turn them into a visual profile of where the class actually stands. Objectives and activities are then set on this evidence rather than on experience alone.

The second step lets the assessment data decide the teaching. Broad objectives are split into targets that build on each other, and the scaffolds differ from child to child—more support for those who had never noticed roots, less for those who already had.

The third step belongs to the children. Tasks mix physical and virtual materials—real bark to touch, macro photographs to examine, root shapes projected on a screen—and call on several senses at once. Children handle the materials, run into problems their current ideas cannot solve, and rebuild their understanding piece by piece.

The fourth step looks at growth rather than outcomes. Process tracking and pre-post comparisons show what has changed in cognition and development, and whatever we learn goes into the next round. The loop closes; the next cycle starts one turn higher.

Pre-course Questionnaire for "The Big Tree and Its Friends"	
Question	Picture
Where is the big tree? Where is the grass? Where is the flower? (The children point them out on the picture.)	
"Where are the leaves? Where are the branches? Where are the roots? Where is the trunk?" (The children point them out on the picture.)	
"You're going to walk from the street lamp over there to the big tree. How will you go? (Will you take the stone path or walk across the grass?)"	

Category	Specific Manifestations	No. of Children
Plant cognition (basic)	Recognize common flowers, grass and trees	12
Plant cognition (advanced)	Recognize several species; begin to notice structure	5
Environmental perception (basic)	Touch and notice textures of stones, bark, etc.	13
Environmental perception (to be refined)	Notice details of texture; expression needs scaffolding	7
Total participants	—	22

Distribution of Survey Data — "The Big Tree and My Friends" (Junior Class)

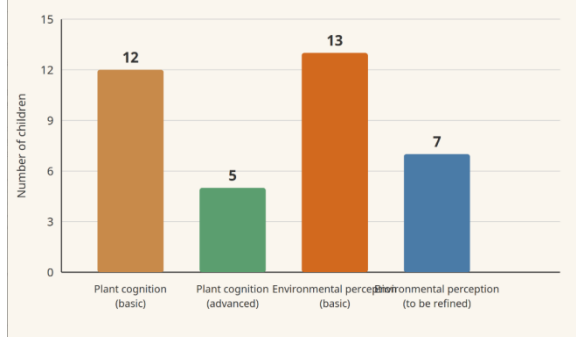


Figure 1. Framework of the three-loop, four-step dynamic assessment model

3. Practical Implementation

The study ran in the junior class (children aged 3–4) of a kindergarten, working with a science theme called "The Big Tree and Its Friends," and the implementation unfolded in phases over two months. Twenty-two children took part, together with their class teacher and one researcher. Figure 2 shows how the four steps played out here; the loop was the same throughout, but the tools changed from stage to stage.

3.1 Phase 1: Pilot Diagnosis—Knowing the Learners Before Teaching Them

Before designing anything, we needed to hear the children out. Each child was taken aside for ten minutes or so with four picture cards—a whole tree, a patch of bark, a tangle of roots, a tree with birds on it—and asked four open questions, among them "What does this tree need to stay alive?" and "Does the tree have friends?" The answers were recorded on a tablet exactly as spoken, transcribed by speech-recognition software, and then read line by line; recurring words such as "milk," "sugar," and "mum" were counted by hand. Reading through the 22 transcripts took less than a week, and it told us three things.

They knew the tree above ground and not below it. 54.5% could name leaves, flowers or fruit, while only 22.7% mentioned roots or the trunk. When asked to draw a tree, most children drew a green cloud on a brown stick. What they left out was precisely what the soil hides, and we started referring to this gap as the "visual blind zone."

Their ideas about how a tree lives came from home. 68.2% said the tree must drink milk, eat sugar, or be fed by its parents; one child put it this way: "Mum gives me milk, so the tree drinks milk too." Only 9.1% said water, and fewer still said sunlight. Animism is expected at this age, but the sheer weight of the milk-and-sugar answer convinced us to confront it head-on rather than hope it would fade.

What they wanted to know was social, not botanical. 45.5% were curious about the tree's "friends" and "family," 31.8% about the look of branches and leaves, and 22.7% about how a tree grows. Not one child asked about roots—which is why, later on, we entered the hidden parts through the "friends" they cared about, the animals living in the bark and around the roots.

These three findings rewrote the objectives. The original aim, "understand trees," was far too loose to teach from, so it was split into three: name the parts of a tree that cannot be seen; find out by experiment what a plant really needs; and notice that the living things around a tree depend on one another.

3.2 Phase 2: Optimized Intervention—Teaching That Follows the Diagnosis

Each of the three aims was then worked on through a mix of real materials and digital support.

For the hidden parts, the children went outside, felt the bark with their hands, and took rubbings

with paper and crayon. Back in the room, the rubbings were photographed on a tablet and the texture pulled up on the screen until the cracks were visible to everyone at once. Roots were harder—nobody can lift a camphor tree—so we washed the soil off a potted seedling and put it under the document camera, where the root hairs showed clearly. Within a week, most children used the word “roots” on their own and pointed to the ground when asked where they were.

The milk-and-sugar belief did not give way so easily. Three seedlings were planted in transparent cups on the windowsill: one watered, one given milk, one left dry. Every two or three days the children photographed their own cup with a tablet and added a mark to a “growth diary” pasted on the wall—a drawing, or a sentence the teacher wrote down for them. By the second week the milk had soured and the dry seedling had drooped, while the watered one had put out new leaves. The point was not to correct them from above; it was to let the cups do the arguing. Three conditions were set up—normal watering with light, water deprivation, and light deprivation. The children cared for the plants themselves, kept observing them, and recorded the growth changes with a documentation application. Comparing the real outcomes side by side, they arrived independently at the conclusion that plant growth requires water and sunlight—an evidence-based explanation taking the place of anthropomorphic fantasy.

At the level of higher-order thinking, the goal was to cultivate ecological awareness. The children watched a science video about woodpeckers and trees and then joined a “tree theater” role-play, taking on the parts of the tree, the pests, and the woodpecker. Acting the relationships out with their bodies, they came to understand biological symbiosis; an abstract ecological notion became concrete experience.

Across the implementation, the teacher’s role moved toward that of context designer, inquiry supporter, and observer of learning. One moment

from the augmented-reality root observation episode stands out: when the three-dimensional root model appeared, the children gasped audibly. That kind of visual impact and embodied experience is something verbal explanation alone cannot produce—and it is exactly the ground on which deep learning is laid.

3.3 Phase 3: Outcome Evaluation—Seeing What Changed

Data collection ran alongside the teaching for the full two months rather than after it. At the end the children were interviewed once more, using the same questions and the same recording procedure. What the comparison shows is reported in the rest of this section. Recognition of concealed structures such as the roots rose from 22.7% to 88.0%, and the share of children who could state the requirements of plant growth in scientific terms went from 9.1% to 90.9%. The core concepts had largely taken hold, and the animistic interpretations had mostly given way. The quality of the children’s inquiry told much the same story. During the process we had collected data on how long the children stayed attentive, how often they interacted, and how they used the tools, and we assembled these into an individual growth portfolio for every child. Leafing through a portfolio, one could actually see how far each child had come—in initiating inquiry on their own, in sharing and working with peers, and in handling the instruments.

The most persuasive evidence of deep learning, however, was qualitative. In the experimental activities, all children (100%) knew that plants must be watered and that unwatered plants would die—and, more tellingly, they took to watering the plants on their own initiative every morning after breakfast. Voluntarily carrying a newly constructed concept into daily routine, with no prompting, suggests that the scientific understanding had begun to consolidate. This is precisely the behavioral signature that distinguishes deep learning from surface recall.

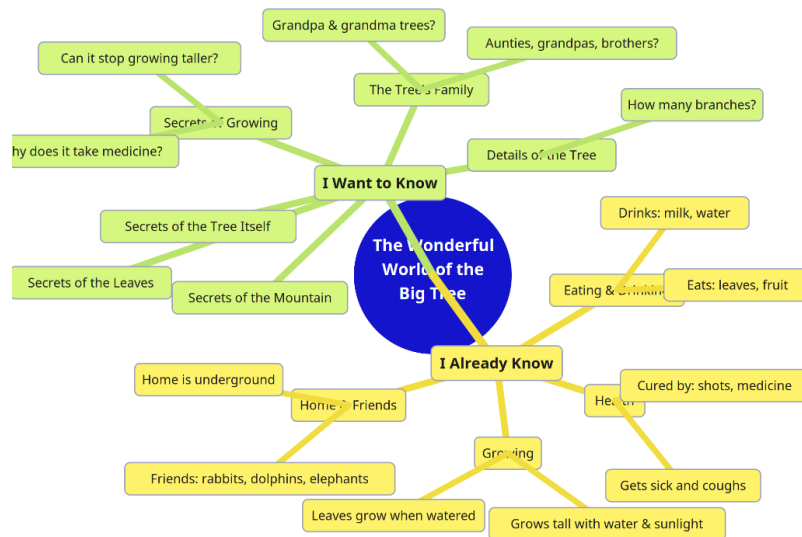


Figure 2. Distribution of digital-intelligence tools across the stages of preschool science education

4. Outcomes and Core Value

The significance of this practice reaches well beyond the optimization of a single science theme. It amounts to a new paradigm for

evaluation reform in preschool education, and it redefines the teacher's role in the era of digital intelligence (see Table 1). Three aspects of its value stand out.

Table 1. Comparison of teacher role transformation in the era of digital intelligence

Dimension	Traditional role	Transformed role in the digital-intelligence era
Instructional design	Experience-based planning driven by intuition and generic templates	Evidence-informed design grounded in diagnostic learner profiles
Curricular positioning	Passive executor of prescribed curricula	Research-oriented practitioner and autonomous decision-maker
Assessment practice	Summative judgment delivered after instruction	Continuous, formative diagnosis accompanying the learning process
Allocation of effort	Repetitive, low-efficiency preparation and manual recording	Deep observation of children and high-quality teacher-child interaction
Basis of decisions	Personal experience and tacit intuition	Data-informed, iterative, evidence-based decision cycles

The first is a paradigm shift in assessment—from terminal judgment to navigational guidance at the point of departure. Under the assessment-first model, conventional assessment of learning is converted into assessment for learning and assessment as learning (Black & Wiliam, 1998; Wiliam, 2011). Assessment stops being an appendage of instruction and becomes the core engine that drives instructional adjustment and supports children's learning—which is what the reform requirement that evaluation empower education looks like when it is substantively enacted.

The second is the restructuring of the instructional model into a data-driven, spiraling closed system. By threading data through diagnosis, design, enactment, and reflection, the closed loop repairs the traditional fragmentation of teaching, learning, and assessment. Teachers, meanwhile, are repositioned—no longer passive executors of curriculum but research-oriented practitioners who make decisions on evidence, with their capacities in curriculum design, learning-state diagnosis, and evaluative reflection strengthened through repeated cycles. Just as important, the model rebuilds teachers'

sense of professional worth. Supported by intelligent assessment tools, teachers are relieved of experience-based, low-efficiency preparation and can give their energy to deep observation of children and high-quality teacher-child interaction. Their identity shifts from passive implementer to data-informed agent of educational research and decision-making, and their professional well-being rises accordingly—grassroots support, in effect, for a teacher evaluation system oriented toward educational outcomes.

The third aspect is the deepening of assessment's educational value, which contributes to a home-kindergarten-community collaborative community. Visualized artifacts—AI-generated growth portraits, learning stories, process documentation—effectively dissolve the information barrier between home and kindergarten. Parents can directly apprehend their children's learning trajectories and conceptual transformations, shifting from passive recipients of educational information to informed supporters of learning. In the survey data gathered during this practice, parental understanding of, and satisfaction with, the kindergarten's education exceeded 95%, and a child-development-centered, data-driven, multi-party ecology integrating teaching, learning, assessment, and upbringing has begun to take shape.

5. Conclusions, Limitations, and Future Directions

What does evaluation owe to the building of a strong education system? From the classroom level, our answer is a modest one: assessment earns its place by diagnosing, guiding, and improving—not by screening and certifying. The assessment-first, three-dimensional closed-loop model described in this article was built with the demands of current evaluation reform in mind, and it grew out of problems we kept meeting in preschool science education—inquiry that stayed shallow, teaching planned from experience, assessment that arrived too late. Because the model is assembled from ordinary tools and ordinary classroom routines, other teachers can pick it up and adapt it to their own settings. If it contributes anything, we would put it this way: by changing what assessment does, it changes the shape of instruction. The question a teacher asks at the end of the day is no longer “what did I teach?” but “what did the children learn, and how did they learn it?” That shift, in our view, is

one workable way for evaluation reform to genuinely take root in preschool education.

Several limitations should be acknowledged, and they delineate directions for future work. First, technological application needs to move from functional assistance toward intelligent integration. Digital tools did improve the precision of assessment, but a risk of “technology for technology's sake” was also observed, potentially entailing over-documentation and additional teacher burden. Future development should pursue more intelligent assessment systems—for example, predictive analytics that anticipate cognitive difficulties or surges of interest from behavioral data such as tool-use frequency and peer-interaction patterns, and intelligent formative support that automatically captures children's peak moments of inquiry, such as posing original questions or completing successful collaboration, and drafts learning-story texts, freeing teachers for higher-value educational interaction. Second, the present assessment concentrated on observable behaviors and explicit concepts; the evaluation of tacit, higher-order dispositions—scientific spirit, inquiry attitudes, and creative thinking—remained insufficiently deep and awaits further breakthrough. Third, the model was piloted only in the domain of science. Whether, and how, it can be adapted to domains such as art and social development, where evaluation criteria are more plural, requires further empirical testing.

Fund Project

This work was supported by the Fifth Batch Project of *Research on Collaborative Innovation in Educating Students through School-Family-Society Cooperation* of the Institute of Curriculum and Textbook Research, Ministry of Education, titled “Research on the Practice of Scientific Activities for Collaborative Innovation in Kindergarten-Family-Society Education Empowered by Digital Intelligence” (Project No. JCS2025036). (教育部课程教材研究所“校家社协同创新育人研究”第五批课题《数智赋能下的园家社协同创新育人科学活动实践研究》项目编号: JCS2025036)

References

- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. Holt, Rinehart and Winston.
- Biggs, J., and Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). Open University Press.

- Black, P., and Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5(1), 7-74.
- Carey, S. (1985). *Conceptual change in childhood*. MIT Press.
- CPC Central Committee and State Council. (2020). *Overall plan for deepening educational evaluation reform in the new era*. People's Publishing House.
- Marton, F., and Saljo, R. (1976). On qualitative differences in learning: I. Outcome and process. *British Journal of Educational Psychology*, 46(1), 4-11.
- Ministry of Education of the People's Republic of China. (2012). *Guidelines for learning and development of children aged 3-6*. Capital Normal University Press.
- National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. National Academies Press.
- Posner, G. J., Strike, K. A., Hewson, P. W., and Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211-227.
- Vosniadou, S., and Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24(4), 535-585.
- Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press.