NOVARA
CONSULTING GROUP

Measured Against Her: AI-Written Goals, Progress Reports and the Deaf Child

Heather Grizzle · November 14, 2026

Research says AI helps teachers write better-formed IEP goals. For a deaf child, a well-formed goal can still be aimed at the wrong target and measured by someone who cannot see it.

Contents

1. The Goal in Part One
2. What the Law Asks of a Goal
3. The Research Says the Tools Write Better Goals
4. Better at What?
5. Four Out of Five
6. The Progress Report Writes Itself
7. Measuring in Her Language
8. Mateo's First Goals, in Three Languages
9. Hannah and the Goal That Was Too Low
10. Eli's Goals
11. The Goal Nobody Wrote for Owen
12. Who Enters the Data
13. What Ambitious Means
14. A Goal Bank Is Not a Goal
15. What a Family Can Do With a Progress Report
16. What Pine Lake Changed
17. What the Tools Could Do Instead
18. Measured Against Her
19. Questions for Governance Review

ABOUT THIS PIECE

This is the tenth part of NCG's AI x IEP series. Part one opened with a goal a generator wrote for Nora, a composite Deaf nine-year-old: that she would produce grammatically correct sentences in four out of five opportunities, as measured by teacher observation.[1] Parts seven through nine followed Pine Lake's deaf and hard of hearing students through the special factors, the screener and the evaluation.[2] This part reaches the center of every IEP, the annual goals, and the progress reports that are supposed to tell a family whether those goals are being met. It ends, as each part now does, with questions for governance review. Pine Lake, its students and its staff are fictional composites, and the goals and reports quoted are illustrative. The research, the federal rules and the National Deaf Education Data Index figures are real.

IN BRIEF

- The annual goals are the part of an IEP the law cares about most and the part AI drafting tools are most often bought to write. Several studies have found that teachers using ChatGPT write goals rated as higher quality, in less time, than teachers working alone.
- Those studies are right about what they measure. The rubrics rate a goal's form: whether it is measurable, specific and functional. They do not ask whether the goal is about the right thing for the child, in her language, measured by someone who can see it.
- A goal can score well on every quality criterion and still be aimed at the wrong target. "Four out of five opportunities, as measured by teacher observation" is a measurable criterion. For a Deaf child taught by a teacher who does not sign, it measures the wrong language through the wrong observer.
- The same tools increasingly draft the quarterly progress report. When software writes the goal, a paraprofessional enters the data and software writes the report on whether the goal was met, the family receives a narrative no person wrote about a target no person chose.
- The Supreme Court requires an IEP to be appropriately ambitious in light of the child's circumstances. A generator anchored on last year's plan and this year's English data produces goals ambitious about the deficit and silent about the child.
- The fix is not to stop using the tools for goals. It is to require that every goal for a deaf or hard of hearing child name its language, its measure and its measurer, and that every progress report show the data, not just the sentence.
- The piece closes with sixteen questions for governance review.

The Goal in Part One

Ten parts ago, this series quoted a goal. A platform at Pine Lake had generated it in under a minute from the data in Nora's file, and her case manager, who knew Nora and had watched her sign, changed a grade level and saved it. It read: by the end of the IEP period, Nora will improve her expressive language skills by producing gram-

matically correct sentences in four out of five opportunities, as measured by teacher observation and work samples.[1]

Part one asked the obvious questions about that goal: grammatically correct sentences in which language, observed by a teacher who may not sign.[1] The series then spent nine parts on everything around it, the contract, the vendor terms, the audit, the records request, the hearing, the special factors, the screener and the evaluation, and each of those parts was, in one way or another, about how that goal came to be written and why no one noticed what it was aimed at. This part comes back to the goal itself, because in March, when Nora's team reconvened after the hearing, they had to write a new one, and the platform was still there, offering to help.

What the team learned writing that goal is the subject of this piece. It turns out that the AI tools that help teachers write IEP goals are, by the measures researchers use, rather good at it. It also turns out that those measures do not ask the question that mattered most for Nora, and that the tools are now writing the progress reports on the goals they wrote.

What the Law Asks of a Goal

The annual goals are the operational heart of an IEP. Federal rules require each IEP to include a statement of measurable annual goals, academic and functional, designed to meet the child's needs that result from her disability, to enable her to be involved in and make progress in the general education curriculum, and to meet each of her other educational needs that result from her disability. The IEP must also describe how the child's progress toward those goals will be measured and when periodic reports on that progress will be provided to the parents, such as quarterly reports issued with report cards.[3] And the team must review the IEP at least once a year and revise it to address any lack of expected progress toward the goals.[4]

In 2017, the Supreme Court gave those requirements a standard. In Endrew F. v. Douglas County School District, it held that an IEP must be reasonably calculated to enable a child to make progress appropriate in light of the child's circumstances, that the plan must be appropriately ambitious in light of those circumstances, and that every child should have the chance to meet challenging objectives.[5] The Court rejected the idea that a plan aimed at merely more than trivial progress was enough. The word it used, ambitious, has become the measure every goal is held to in a dispute.

Read together, those rules ask four things of a goal. It has to target a need that comes from the child's disability. It has to be measurable. Someone has to actually measure it, on a schedule, and report what they find to the family. And it has to be ambitious for this child, not for the average child with her eligibility label. A goal can satisfy the second requirement perfectly and fail the other three.

The Research Says the Tools Write Better Goals

It would be dishonest to write this piece without starting with the evidence for the tools, because it is better than many critics assume.

The most cited study is by Salih Rakap, published in the Journal of Special Education Technology in 2024. He randomly assigned twenty-two novice special education teachers either to write IEP goals for five children with autism using ChatGPT or to write them without it, and had the goals rated on a published goal-quality instrument. The teachers who used ChatGPT wrote goals rated significantly higher in quality, more comprehensive and more specific to each child's strengths and needs, and spent significantly less time doing it, regardless of whether they had prior training in goal writing.[6] A smaller mixed-methods study of special educators in inclusive primary classrooms in Greece, published in Education Sciences in 2025, found that ChatGPT contributed to clearer goal-setting and more structured planning, while the teachers stressed the need for critical oversight and for protecting the relational side of teaching.[7] And a University of Virginia study described in August 2026 found that ChatGPT could improve the quality of learning goals written by teachers, and that the goals it wrote were of comparable quality to those written by more experienced teachers; its researcher, Danielle Waterfield, said the tool produced comparable quality in creating IEP goals.[8] National reporting earlier in 2026 summarized research from Virginia and the University of Central Florida the same way: used appropriately, AI can help special education teachers write IEPs of equal or higher quality than they produce alone.[9]

That is a consistent finding across different designs and countries, and it fits what teachers report. Goal writing is hard, the standards for a well-formed goal are specific, and a language model trained on many goals is good at producing sentences that meet those standards. A novice teacher who has never written a measurable goal for a child with autism gets, from the tool, a model of what one looks like. That is a real benefit,

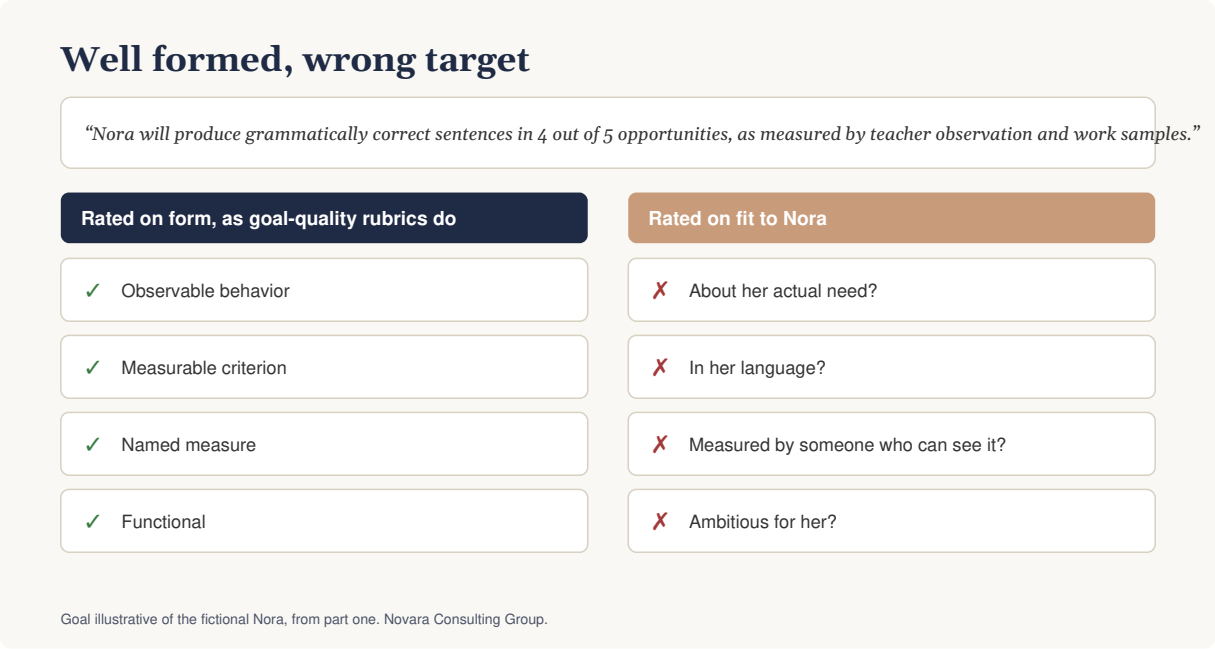
and in a profession where forty-five states reported shortages of special education teachers in the 2024 to 2025 school year, it is not a small one.[9]

Better at What?

The question the research leaves open is what "quality" means in those studies, and the answer is specific. Goal-quality instruments of the kind these studies use rate a goal's form: whether it describes an observable behavior, whether it states a measurable criterion, whether it is functional, meaning useful in the child's everyday life, whether it is set in a meaningful instructional context. Those are real virtues. A goal that has them is easier to teach to and easier to measure than one that does not.

What those instruments cannot rate is whether the goal is aimed at the right target for the particular child, because the rater usually does not know the child. They read the goal and the case description they were given. If the case description says a child has significant delays in expressive language, a goal targeting expressive language is well aimed by the instrument's standards. If the case description is wrong, as Nora's was for three years, the instrument has no way to know.

Nora's original goal makes the point. Rate it on form and it does reasonably well. It names an observable behavior, producing sentences. It states a criterion, four out of five opportunities. It names a measure, teacher observation and work samples. It is arguably functional, since writing grammatical sentences is useful. A rater who had never met Nora could give it a respectable score. And it was aimed at the wrong thing in nearly every respect that mattered: it treated a gap in a second language as a delay in language, measured it in English through a teacher who did not sign, and never mentioned the language Nora actually commands.



Nora's generated goal, rated two ways. On form, it passes. On fit to the child, it fails every question that mattered.
Novara Consulting Group; goal illustrative of the fictional Nora

This is not a flaw in the studies. They asked whether the tools help teachers write well-formed goals, and the answer is yes. It is a limit on what their answer can be used to claim. A district that adopts a goal generator because research shows it improves goal quality has evidence that its goals will be better formed. It has no evidence that they will be better aimed, and for a deaf child, aim is most of what matters.

Four Out of Five

Look at enough generated goals and a pattern emerges in the criteria. Four out of five opportunities. Eighty percent accuracy. Three consecutive data collection sessions. As measured by teacher observation, work samples and data collection. These phrases are not wrong. They are the most common criteria in the goals the models were trained on, and so they are the criteria the models produce most often. They are also, for many children, close to meaningless.

A criterion means something only if someone can apply it. "Four out of five opportunities" assumes that the opportunities are defined, that someone counts them, and that the person counting can tell a success from a failure. For Nora's original goal, the opportunities were sentences, the counter was a general education teacher who does not

sign, and the successes were English sentences on paper, since the teacher had no way to judge a sentence in ASL. The criterion was measurable in form and unmeasured in practice. Nobody counted. The progress reports, as the next section describes, were written without the count.

"As measured by teacher observation" carries the same assumption more quietly. It names a teacher without saying which one, and for a deaf child the answer matters more than the method. A teacher of the deaf who signs fluently, observing Nora explaining a science concept in ASL, can tell whether her language is growing. A general education teacher who does not sign, observing the same moment through an interpreter, sees the interpreter's English. A goal that says "teacher observation" without saying which teacher, in what language, has not said how it will be measured. It has said that someone will look.

When Nora's team rewrote her goals in March, they wrote the measurer into each one. Her ASL narrative goal names the teacher of the deaf and a video sample collected each quarter and scored against a narrative rubric. Her English writing goal names the teacher of the deaf and the general education teacher together, scoring her written work in English with a rubric built for deaf bilingual writers, which counts what she is learning in English as a second language rather than every error a native speaker would not make. Neither goal says four out of five.

The Progress Report Writes Itself

The criterion was measurable in form and unmeasured in practice.

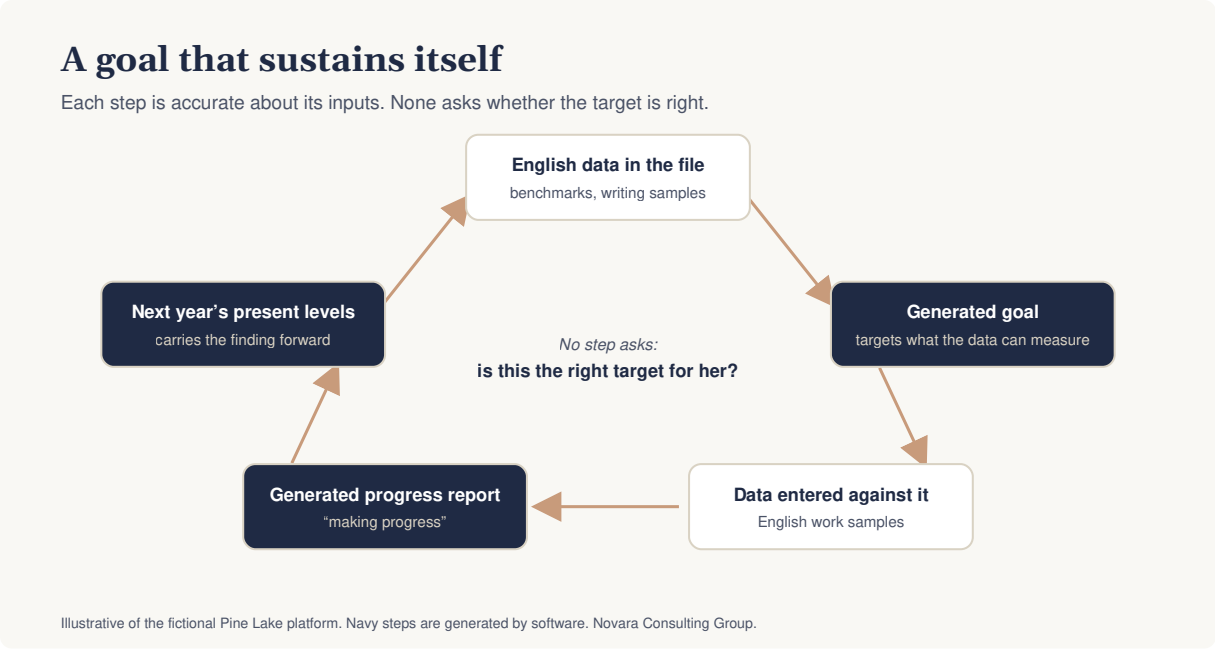
The goal is half of the problem. The other half arrives four times a year, and it is newer.

Federal rules require the IEP to say when periodic reports on a child's progress will be provided to the family.[3] At most districts that means quarterly progress reports, issued with report cards, and each one states, for each goal, whether the child is making progress and how much. Writing those reports is one of the most time-consuming tasks in a special education teacher's year: a case manager with twenty-six students and four goals each writes more than four hundred progress statements a year. IEP platforms increasingly offer to draft them, and the drafts are built from whatever progress data the platform holds.

Pine Lake's platform added the feature the same year as goal drafting. A case manager or a paraprofessional enters data points against each goal through the quarter. At the end of the quarter, the platform drafts a progress statement from the goal and the data: "Nora is making progress toward this goal. She has demonstrated improvement in producing grammatically correct sentences and will continue to work on this skill." The case manager reviews it, edits it or not, and the report goes home.

For Nora's original goal, the data points came from her English work samples, the only thing anyone was counting. Her English writing did improve over the year, slowly, as it does for a bilingual deaf child learning to write a second language. So the generated report said she was making progress, which was true of her English and said nothing at all about the language delay the goal claimed to target, because there was no such delay to measure. The report was accurate about the data and wrong about the child, the same failure as the goal, now repeated quarterly and sent home to her mother.

There is a further step, and it closes a loop. When the next annual review came, the IEP generator read last year's IEP, including the goals, and last year's progress reports, including the generated statements that Nora was making progress on her expressive language goal. It produced a new present levels statement that carried the language delay forward, noted progress, and proposed a new goal at a slightly higher criterion. Part four found the same sentence about significant delays in expressive language carried forward for three years.[10] This is the machinery that carried it. A machine wrote the goal, a person entered English data against it, a machine wrote the report on it, and a machine read the report as evidence for the next goal. At no step did anyone ask whether the goal was about the right thing, because at no step was that the question the step was built to answer.



How a goal no one chose can sustain itself. Each step is accurate about its inputs; none asks whether the target is right.
Novara Consulting Group; illustrative of the fictional Pine Lake platform

Measuring in Her Language

The alternative is not complicated to describe and is hard to staff. A goal for a deaf child who signs, aimed at her language, has to be measured by someone who can see her language. For young children, standardized tools exist: Gallaudet University's VL2 center publishes the Visual Communication and Sign Language Checklist for children from birth to five, and schools for the deaf use a range of receptive and expressive ASL measures.[11] For older children, the most common method is the simplest: a video sample of the child signing, collected on a schedule, scored by a fluent signer against a rubric, and kept so that the next sample can be compared with the last. That is more work than a tally of English sentences, and it is the only way to know whether the goal is being met.

The states that have tried to measure deaf children's language at scale show how far most systems are from that. Iowa's 2022 language milestones law requires language assessments for deaf and hard of hearing children from birth to age eight, and the state's own report, as the National Deaf Education Data Index records, shows that 156 of 1,824 children, about nine percent, were assessed, more than half of them below average.[12] California has published an annual report on young deaf children's language

development, required by its SB 210, every year since the 2016 to 2017 school year.[12] Those are the leading states. In most of the others, no one is measuring a deaf child's language in her language at all, and a goal that targets it will be measured, if it is measured, in the only data the system has, which is English.

That is the context in which a goal generator operates. It does not invent the English bias in a deaf child's file. It inherits it, and then it formalizes it, turning whatever the file can measure into goals with criteria that fit the measurement. The research summarized in part eight is worth recalling here: for deaf readers, the strongest predictor of reading achievement is language ability in whatever language the child has, not phonological skill in English.[13] A goal system that can only measure English, and therefore only writes goals about English, is aimed at the weaker lever for the children who most need the stronger one.

Mateo's First Goals, in Three Languages

Mateo, the five-year-old from parts seven and eight who signs fluently, is learning to read English through ASL and lives with parents who speak Spanish, had his first full set of kindergarten goals written the same spring.[2] The platform drafted five, all in English language arts and all keyed to the kindergarten standards: letter-sound correspondence, phonological awareness, high-frequency words, concepts of print and oral language. Every one was well formed. Two of them, phonological awareness and oral language, were aimed at skills that depend on hearing spoken English and that part eight showed the screener could not measure in him.[14]

His team kept concepts of print, which is about how books and text work and transfers directly. They rewrote letter-sound correspondence as letter-fingerspelling-print correspondence, the way deaf children are commonly taught to connect the alphabet to English words, and named his teacher of the deaf as the measurer. They replaced oral language with an ASL language goal, measured by a quarterly video sample. And they added one goal the platform had no way to propose, because nothing in Mateo's school data described his home: that his parents would receive every progress report in Spanish, with a fifteen-minute meeting each quarter through a Spanish interpreter and an ASL interpreter, so that the people who know him best outside school could see what the people who teach him were measuring. A progress report a family cannot read is not a report to the family, whatever the regulation's schedule says.

Hannah and the Goal That Was Too Low

The risk for some deaf and hard of hearing children is not that the generated goals are aimed at the wrong deficit. It is that the child is doing well enough that the generator, and the team, aim at almost nothing.

Hannah, the eight-year-old with cochlear implants who reads above grade level, received two generated goals: one to maintain her reading comprehension at grade level and one to use her hearing technology consistently.[2] Both were well formed. Her parents found them insulting, and they had history on their side. The first special education case the Supreme Court ever decided, Board of Education v. Rowley in 1982, concerned a deaf girl, Amy Rowley, who was doing well in a regular classroom without a sign language interpreter. Her school argued that because she was passing from grade to grade, she was receiving an appropriate education, and the Court agreed, holding that the law did not require a school to maximize a child's potential.[15] For thirty-five years, Rowley was read by many districts as permission to aim low for any child who was keeping up. Endrew F. corrected that reading in 2017, holding that a child's program must be appropriately ambitious in light of her circumstances, for a child performing at grade level as much as for one who is not.[5]

What would an ambitious goal for Hannah look like? Her parents had an answer. Hannah understands her teacher well in a quiet room and misses most of what her classmates say in group work and on the playground. She has no deaf or hard of hearing friends at school. Her reading is strong, but her vocabulary for the abstract words she will meet in fourth and fifth grade science and social studies is behind her decoding, because so much of that vocabulary is learned incidentally, by overhearing, which she cannot do. Her team wrote goals for vocabulary in the content areas, for strategies to follow and join group discussion, and for regular contact with other children who use hearing technology. None of those needs appeared in her data as a deficit, because her scores were good. Every one of them comes from her hearing loss, and every one of them is the kind of need the law requires an IEP to address.

Eli's Goals

Eli, the sixteen-year-old from part seven who is Deaf, autistic and largely nonspeaking, makes the same point from a different direction.[2] Eli communicates through a combination of ASL signs, a speech-generating device and pictures. His team has writ-

ten communication goals for years, and they are among the most carefully built in the district: which signs he uses spontaneously, which he uses with prompting, how many device messages he initiates in a day, in which settings, with which partners.

When Pine Lake's platform began drafting goals, it drafted Eli's from his file. It proposed a goal that he would increase his expressive vocabulary to fifty words, measured by speech-language pathologist observation, and a goal that he would answer "wh" questions with eighty percent accuracy. Both are standard goals for a nonspeaking autistic student. Neither said in which mode. A speech-language pathologist new to Eli, reading those goals in September, could reasonably count spoken words, of which Eli has few, or device words, or signs, and the progress reports would describe three different children depending on which she chose.

The federal rules require Eli's team to consider whether he needs assistive technology devices and services, and the special factors require them to consider his language and communication needs as a deaf child.[16] His real goals do both, because the people who wrote them knew him. The generated goals did neither, because a goal that names no mode is easier to generate and reads as complete. His team kept the generated wording for one goal, about requesting breaks, because it was good, and rewrote the rest. The time the tool saved on that one goal was real. So was the time it cost to notice that the others had dropped the most important fact about him.

The Goal Nobody Wrote for Owen

Sometimes the problem with a generated set of goals is not a goal that is wrong but a goal that is missing. Owen, the hard of hearing seventh grader who uses spoken English and hearing aids, had three goals on his generated draft: one in reading comprehension, one in written expression, and one in organization.[2] All three were reasonable and supported by his data.

What his mother asked for, and what his team added in the spring, was a goal no data in his file would have suggested: that Owen would advocate for his own access. That means telling a new teacher how to use his remote microphone, asking a classmate to repeat what he missed, checking that captions are on before a video starts, and telling an adult when his equipment is not working. For a hard of hearing teenager in a hearing school, those skills decide whether every accommodation in his plan actually happens on a given day, and they are among the skills he will need most after he leaves

school. They are also skills that appear in his file only by their absence, in the cafeteria incident part eight described, in the afternoons he fell behind, and in the days the microphone sat in a drawer.[14]

A generator proposes goals for the needs that the data show. It cannot propose a goal for a need the data show only as a pattern of failures attributed to something else. That is one more reason the goals for a deaf or hard of hearing child need a person who knows deaf and hard of hearing children, and not only a person who reviews the draft.

Who Enters the Data

A progress report is only as good as the data entered against the goal, and for deaf and hard of hearing students, who enters that data is a question with consequences. At Pine Lake, as at many districts, progress data for some goals were entered by paraprofessionals, and in two cases by an educational interpreter who was asked, because she was in the classroom all day, to record whether the student met the criterion in each opportunity.

That second arrangement deserves a pause. An educational interpreter's job is to convey communication between a deaf student and the people around her, accurately and without adding to it. Asking her to judge, in the same moment, whether the student's work meets a goal puts her in the position of evaluating the communication she is supposed to be carrying, and it makes the quality of the interpretation part of the measurement without anyone recording that it is. A goal that names "teacher observation" and is in practice measured by the interpreter has not been measured by the person the IEP says.

Generated progress reports make that problem harder to see, not easier. A report drafted from data points does not record who entered them. It produces a fluent sentence about progress regardless of whether the data came from a teacher of the deaf scoring an ASL video sample against a rubric or from an interpreter ticking a box between sentences. Pine Lake's third rule, that every progress statement show its data, the dates and the collector, was written with that in mind. A family reading that a goal was measured by the educational interpreter now knows to ask a question it could not ask before.

What Ambitious Means

The Supreme Court's word in Endrew F. was ambitious, and it was chosen to raise expectations.[5] It is worth asking what an AI goal generator does to ambition, because the answer is structural, not a matter of any one product.

A generator writes the most likely goal given the inputs. The inputs are last year's IEP, this year's data, and the eligibility category. The most likely goal given those inputs is a modest step up from last year's goal, at a common criterion, in an area the data show as a weakness. That is a reasonable default, and it is the opposite of ambition in the sense the Court meant. Ambition, for a particular child, usually means a goal that the data do not suggest, because the data were collected under conditions that hid what the child could do. For Nora, it meant a goal for ASL narrative at a level appropriate for a fluent nine-year-old signer, which no one at Pine Lake had measured before. For Owen, it meant a goal for self-advocacy that his data recorded only as problems. For Eli, it meant a goal for spontaneous signed requests to peers, not only adults, which no one had tried.

A deficit-anchored data file produces deficit-anchored goals. Deaf education has a long history of low expectations built on that kind of file, in which a child's English scores stood in for her intelligence and her language. The generator did not create that history. It is very good at continuing it, in polished, measurable, correctly formatted sentences, unless someone at the table asks what this child could do if the plan were built around what she already has.

A Goal Bank Is Not a Goal

None of this is new in kind. Long before any language model, districts maintained goal banks: libraries of pre-written goals organized by area and grade, from which a case manager could pick and adapt. Goal banks had the same strengths and the same risks as generators. They saved time, modeled good form and spread common criteria across a district. They also produced plans in which the goals for different children with the same label were nearly identical, and in which the most common goals in the bank became the most common goals in the district whether or not they fit.

An AI goal generator is, in one sense, an infinitely large goal bank that adapts its entries to the data it is given. That adaptation is the improvement, and it is real: the goals it produces mention the child's name, her grade, her scores and her strengths,

and so they look tailored in a way a bank entry never did. That appearance is also the new risk. A case manager picking a goal from a bank knew she was choosing a template and adapting it. A case manager reviewing a generated goal that already uses the child's name and cites her benchmark score is reviewing something that looks like it was written about the child, and is less likely to ask whether it was.

What a Family Can Do With a Progress Report

The progress report is the one document in this process that reaches the family on a schedule without anyone having to ask. That makes it the family's best early warning, if they know how to read it.

A family can treat each quarterly report as a question rather than an answer. If the report says a child is making progress, the family can ask, in writing, for the data the statement is based on: the data points, the dates and who collected them. Those data are part of the child's education records, and part five described the family's right to inspect and review them.[17] If the report says the child is not making expected progress, or if the data show the same level quarter after quarter, the family can ask for an IEP meeting to revise the goal, because the team is required to revise the IEP to address any lack of expected progress.[4] And if the report describes progress on a goal the family believes is aimed at the wrong thing, as Nora's mother came to believe, the report itself is the evidence that the goal needs to change: a year of measured progress on the wrong target is a year in which the right target was not measured at all.

None of that requires a lawyer or a hearing. It requires a family to know that a progress statement without data is not the end of the conversation, and that the regulation gives them the right to see what is beneath it.

What Pine Lake Changed

Dana took four rules to her board in June, after Nora's, Eli's and Owen's teams had each rewritten their goals. None of them restricted the platform. Each of them changed what a goal and a progress report had to contain.

The first was that every annual goal for a deaf or hard of hearing student name three things: the language or mode in which the skill is performed, the measure used, and the person who will collect the measure, by role. A goal that says "teacher observation" without saying which teacher, in what language, would not be accepted. The sec-



ond was that the goals targeting language and communication for those students be written by the teacher of the deaf with the family, not generated, which extended the rule part seven described for the special factors.[2] The third was that every progress report, generated or not, include the data points it was based on, the dates they were collected and who collected them, so that a family could see what the progress statement rested on. A statement that a child is making progress, with no data beneath it, would no longer go home. The fourth was that any goal carried forward from a prior year at the same target be flagged automatically by the platform for the case manager, with the number of years it had appeared.

Pine Lake's four rules for goals and progress reports, illustrative

Rule	What it requires	What it would have caught	
Name the language	measure and measurer	Every goal for a deaf or hard of hearing student	Nora's goal measured in English by a non-signing teacher
Language goals written by the teacher of the deaf with the family	Not generated	Eli's mode-less communication goals	
Progress reports show their data	Data points	dates and collector under every statement	Reports of progress on a target no one measured
Flag carried-forward goals	Platform marks repeats and counts the years	Nora's language delay goal in its third year	

Source: NCG; drawn from 34 CFR 300.320 and 300.324 and the Pine Lake cases

The third rule drew the most resistance, from case managers, who saw it as more work. It is more work. It is also the only way a family can tell whether a progress report describes their child or describes the generator's default sentence for a goal with some data under it. The vendor, asked, said the platform already held the data points and could print them under each statement at no cost. It had simply never been asked to.

What the Tools Could Do Instead

It would be easy to end a piece like this with a list of things teachers should do differently, and Pine Lake's rules are such a list. It is worth asking what the tools themselves could do, because the vendors are better placed than any district to change what every goal looks like.

A goal generator could ask before it writes. For any student whose eligibility, evaluation or special factors indicate a hearing loss, it could require three inputs before drafting a communication or language goal: the language or mode in which the student performs the skill, the measure the team intends to use, and the role of the person who will collect it. It could refuse to complete "as measured by teacher observation" without a role attached. It could offer measures matched to the mode, a video sample and rubric for a signed skill, a device log for an augmentative communication goal, a listening task in noise for a hard of hearing student, instead of defaulting to four out of five opportunities. None of that is technically difficult. It is a matter of which questions the product asks.

A progress report generator could do even more with less. It already holds the data. It could print them under every statement, as Pine Lake's vendor admitted it could. It could decline to write "making progress" when the data show no change across two quarters, and write instead that the data do not show progress and that the team should consider whether to revise the goal, which is what the federal rule requires the team to do.[4] It could flag a goal that has been carried forward at the same target, with the number of years. Each of those features would make the platform's output less fluent and more honest, and each would give a family something real to read.

Vendors will reasonably say that districts have not asked for any of this. Part three of this series found the same answer in the vendors' contracts: the documents say what buyers asked for, and buyers have not asked for a record of what the machine wrote

and what a person did with it.[18] The same is true here. The cheapest way to change every generated goal in a state is for the state, or a large district, or a consortium of districts, to write these questions into the next procurement.

Measured Against Her

Ten parts in, the series has followed one goal from its first generated draft to its rewrite, and the lesson it teaches is narrower than the one critics of these tools usually draw. The tools write well-formed goals. The research says so, and the research is right. What they cannot do is know what a goal should be about, for a child whose file does not describe her, and they cannot measure whether it was met. Those are the two things a goal is for.

Nora's new goals are less tidy than her old one. They name a teacher, a language, a video sample, a rubric and a schedule, and they set a target in ASL that no one at Pine Lake had measured before, so the first quarter's report was mostly a baseline. Her mother says it is the first progress report she has received that sounds like her daughter. The record is not the child, a goal is not a target until someone has asked what the child should be aiming at, and a progress report that shows no data is not a measure of anything.

Questions for Governance Review

These questions are written for school boards, special education directors, state education agencies, vendors, teacher preparation programs and families. They are meant to be asked in a meeting with real goals and real progress reports on the table.

> **FOR SCHOOL BOARDS AND SUPERINTENDENTS**
>
> 1. What share of the annual goals in our IEPs last year were drafted by software, and in which areas? Could we find out if we had to?
> 2. When we evaluate whether our IEPs meet the Endrew F. standard, do we look at whether goals are well formed, or whether they are aimed at the right targets for the child? Who in the district is positioned to judge the second?
> 3. Do our progress reports to families show the data they are based on? If not, what is a family supposed to rely on?

FOR SPECIAL EDUCATION DIRECTORS

4. For each deaf or hard of hearing student, does every goal name the language or mode, the measure and the person who measures it? How many do not?
5. How many goals in our district have been carried forward at the same target for two or more years? Who reviews them?
6. Who measures progress on language goals for students who sign, and in what language is that measurement done?
7. When the platform drafts a progress report, who is accountable for its accuracy: the person who entered the data, the person who reviewed the draft, or the person who signed the report card?

FOR STATE EDUCATION AGENCIES

8. Do state monitoring reviews of IEPs check whether goals for deaf and hard of hearing students name the language of measurement?
9. Does the state measure the language development of deaf and hard of hearing children in their own language, and does that data reach the IEP teams that write their goals?
10. Should progress reports generated by software be required to include the underlying data, as a condition of the platform's approval or use?

FOR VENDORS

11. When your tool drafts a goal for a child whose eligibility category is hearing impairment or deafness, does it ask for the child's language and mode before it writes? If not, would you add the question?
12. Can your platform print the data points under every generated progress statement, and flag goals carried forward at the same target?
13. On what goals was your model trained, and how many were written for deaf or hard of hearing children who sign?

FOR TEACHER PREPARATION PROGRAMS

14. When we teach new special educators to write goals, do we teach them to evaluate a generated goal's fit to the child, not only its form?

FOR FAMILIES AND ADVOCATES

15. For each of my child's goals, in what language will she perform the skill, how will it be measured, and who will measure it?
16. When I receive a progress report, can I see the data it is based on? If my child's goal has appeared before, for how many years, and why has it not been met?

NOT LEGAL ADVICE

This piece describes federal special education rules, published research and how NCG reads them as applied to AI-drafted goals and progress reports. Pine Lake, Nora, Eli, Owen, their families and staff, and the goals and reports quoted are fictional and illustrative; no statement about Pine Lake's platform describes any named product. Families should contact their state's Parent Training and Information Center or an attorney, and districts their own counsel. NCG does not evaluate, endorse or rank any IEP product.

☐ For each goal, ask: is this about my child's real need, or about what the data could measure?

☐ Check that each goal names the language or mode in which my child performs the skill.

☐ Check that each goal names who will measure progress, and that the person can see the skill in my child's language.

☐ Ask whether any goal has been carried forward from a previous year, and why.

☐ Ask that progress reports include the data points, dates and collector, not only a sentence.

☐ If your child signs, ask how her language in ASL is measured and by whom.

☐ Ask whether any goal or progress statement was drafted by software, and who reviewed it.

TIPS

Know something about a sign language AI system that the public should know? Send it to tips@novaracg.com. Every tip is read by me and kept confidential. If you work in the UK, writing to us is not by itself a protected disclosure under the Public Interest Disclosure Act 1998, so before you share employer information, check your options with Protect (protect-advice.org.uk), the free whistleblowing advice line.

SUPPORT NCG

NCG does this work without vendor money, so our findings stay ours. If it has been useful to you, you can support NCG's independent work here. NCG is a for-profit company, not a charity. Contributions are not tax-deductible and buy no service or influence.

Notes

1. Heather M. Grizzle, "Drafted by AI, Signed by a Human," Novara Consulting Group, 15 October 2026, part one of the AI x IEP series. https://www.novaracg.com/2026/10/15/drafted-by-ai-signed-by-a-human/
2. Heather M. Grizzle, "Six Children, One Sentence: The Special Factors a Generator Cannot Write," Novara Consulting Group, November 2026, part seven of the AI x IEP series. https://www.novaracg.com/?p=1438
3. 34 CFR 300.320(a)(2)(i) (measurable annual goals, academic and functional, designed to meet needs resulting from the disability) and 300.320(a)(3) (how progress will be measured and when periodic reports will be provided). https://www.ecfr.gov/current/title-34/section-300.320
4. 34 CFR 300.324(b)(1) (review at least annually and revise to address any lack of expected progress toward the annual goals). https://www.ecfr.gov/current/title-34/section-300.324
5. Endrew F. v. Douglas County School District RE-1, 580 U.S. 386 (2017).

6. Salih Rakap, "Chatting with GPT: Enhancing Individualized Education Program Goal Development for Novice Special Education Teachers," Journal of Special Education Technology 39, no. 3 (2024): 339-348, doi: 10.1177/01626434231211295.
7. Stergiani Giaouri and Maria Charisi, "Enhancing IEP Design in Inclusive Primary Settings Through ChatGPT: A Mixed-Methods Study with Special Educators," Education Sciences 15, no. 8 (2025): 1065, doi:10.3390/educsci15081065.
8. Dakota Griffin, "UVA study explores the role of AI in special education," WVIR/WHSV, 19 August 2026, describing research by Danielle Waterfield, University of Virginia School of Education and Human Development. https://www.whsv.com/2026/08/19/uva-study-explores-role-ai-special-education/
9. Jonaki Mehta, "Overworked and understaffed: Special ed teachers turn to AI for help," NPR, 20 May 2026, citing research by Olivia Coleman (University of Central Florida) and Danielle Waterfield (University of Virginia) and Learning Policy Institute shortage data. https://www.nhpr.org/2026-05-20/overworked-and-understaffed-special-ed-teachers-turn-to-ai-for-help
10. Heather M. Grizzle, "Twenty Files: Auditing the AI-Drafted IEP," Novara Consulting Group, 7 November 2026, part four of the AI x IEP series. https://www.novaracg.com/?p=1423
11. Visual Language and Visual Learning (VL2), Gallaudet University, "Visual Communication and Sign Language Checklist." https://vl2.gallaudet.edu/visual-communication-and-sign-language; Kansas School for the Deaf, "American Sign Language (ASL) Assessments." https://ksdeaf.org/asl-assessments
12. Heather M. Grizzle, "Thirteen States Are Public. Sixty-Four Percent of Their School Units Say Nothing," Novara Consulting Group, 1 November 2026, reporting the National Deaf Education Data Index version 2.0 state files for Iowa and California. https://www.novaracg.com/2026/11/01/thirteen-states-are-public-sixty-four-percent-of-their-school-units-say-nothing/
13. Rachel I. Mayberry, Alex A. del Giudice and Amy M. Lieberman, "Reading Achievement in Relation to Phonological Coding and Awareness in Deaf Readers: A Meta-analysis," Journal of Deaf Studies and Deaf Education 16, no. 2 (2011): 164-188.

14. Heather M. Grizzle, "The Screener Was Listening: AI, Child Find and the Children It Cannot Hear," Novara Consulting Group, November 2026, part eight of the AI x IEP series. https://www.novaracg.com/?p=1442
15. Board of Education of the Hendrick Hudson Central School District v. Rowley, 458 U.S. 176 (1982).
16. 34 CFR 300.324(a)(2)(iv)-(v) (language and communication needs of a deaf or hard of hearing child; whether the child needs assistive technology devices and services) and 300.105 (assistive technology). https://www.ecfr.gov/current/title-34/section-300.324
17. Heather M. Grizzle, "Forty-Five Days: When a Parent Asks for the AI's Draft," Novara Consulting Group, 8 November 2026, part five of the AI x IEP series. https://www.novaracg.com/?p=1429
18. Heather M. Grizzle, "We Read the Terms for the IEP Generators," Novara Consulting Group, 5 November 2026, part three of the AI x IEP series. https://www.novaracg.com/2026/11/05/we-read-the-terms-for-the-iep-generators/

How to cite

Grizzle, H. M. (2026, November 14). Measured Against Her: AI-Written Goals, Progress Reports and the Deaf Child. Novara Consulting Group. https://www.novaracg.com/2026/11/14/measured-against-her-ai-written-goals-progress-reports-and-the-deaf-child/