ENGELMANN-BECKER CORPORATION

Dear Owen: What Does It Mean to "Field-Test" a DI Program?

At Engelmann-Becker Corp, we're often asked what truly differentiates Direct Instruction programs from everything else on the market — including programs that claim to be "explicit," "structured," or "aligned with the Science of Reading."

Most people assume the answer lies in the scripting, the examples, the correction procedures, or the placement tests. Those features matter, of course, but they are *not* what makes Direct Instruction (**DI—big D, big I)** unique.

Recently, Senior DI Author, Owen Engelmann received an email asking a version of the essential question:

> *"What exactly is field-testing, and why is it the keystone that yields consistently successful outcomes separating DI from all other curricula?"*

In addition to responding privately, Engelmann and fellow author Evan Haney prepared the overview below so we can share it more broadly with *anyone interested in what it means to field-test a Direct Instruction program.*

What "Field-Testing" Really Means in DI

Owen Engelmann, Owen.Engelmann@ebc-ed.com
Evan Haney, Evan.Haney@ebc-ed.com

When Siegfried Engelmann and Carl Bereiter began developing Direct Instruction in the 1960s, they operated under a principle that still defines DI today:

If students are not learning as intended, the fault lies in the instruction — and it is the Author's responsibility to fix the instruction.

To fix anything, you need valid, complete, real-world performance data. And to get that data, you must observe full programs taught to real students, with teachers delivering lessons with enough fidelity that you can tell whether the *program* caused the success or the problems.

That is the essence of DI field-testing.
It is **not a pilot.** It is **not a preview.**
It is the engineering process that ensures DI yields successful results.

Two Goals — Often in Conflict

Every DI development and revision cycle must satisfy two goals, and these goals rarely align peacefully:

1. Engineer the most effective instruction possible — increasing clarity, efficiency, generalization, and student engagement.
2. Satisfy the content and pedagogical requirements imposed by states, districts, and adoption agencies — requirements that often have little connection to instructional effectiveness and sometimes directly conflict with it. These requirements often omit or mischaracterize essential skills, include or over-emphasize skills that are inessential or harmful, and specify essential skills but require a sequence that is ineffective, problematic or non-existent.

Field-testing is how we protect instructional quality when external pressures are misaligned with student learning.

ENGELMANN-BECKER CORPORATION

The Five Non-Negotiable Conditions for Valid Field-Test Data

To identify and correct genuine instructional problems, field-testing must meet these conditions:

1. Fidelity to DI classroom protocols

Conscientious grouping, scheduling, seating, management, and teacher preparation are essential. Without these, performance data becomes noise.

2. Presentation close enough to the script to diagnose causes

We must be able to tell whether a performance problem came from:

- the **program design**,
- the **teacher's delivery,**
- the **student's prior performance history.**
- or **environmental influences** (fire drill, announcements, behavior management issues, etc. - this happens A LOT)

This is why, when needed, we provide substantial professional development for field-test teachers — not for compliance, but for receiving the highest quality data and diagnostic clarity.

3. Full recordings of *every* lesson for *every* group for *every* iteration

Audio at minimum; video when possible.
Partial samples rarely reveal root causes.
DI programs orchestrate every prerequisite and application task. If outside programs interfere with that sequence or teachers skip a component, student performance will decline — even if DI lessons are perfectly designed.

4. Timely transfer of recordings to the authors

If early errors aren't addressed quickly, students develop chronic misconceptions. Once those errors spread across lessons, all subsequent data from that classroom may become contaminated and unusable.

5. Accurate attendance and exposure data

A program cannot be faulted for a student who was absent for crucial instruction.

These five conditions exist for one reason: to determine whether the program, as engineered, produces the intended learning.

Critical to note: When revisions introduce substantial changes to wording, sequencing, examples, or skill integration, the program must undergo another iteration of the full field-test cycle. DI programs are not considered internally-validated unless the complete, revised version has demonstrated the intended performance outcomes in actual classrooms.

What Authors Look for When Analyzing Field-Test Data

Each lesson is analyzed for:

- Time required for each task
- Moments where teachers struggle
- Student-response patterns
    - no response
    - delayed response
    - **reasonable incorrect responses** (often the most diagnostically valuable)
    - other errors
    - affective or motivational reactions

A deep understanding of successfully teaching DI is required to interpret these performance problems correctly. A deep understanding of DI engineering is required to efficiently and effectively fix performance problems. **It is easy to see a symptom; it is much harder to identify and cure the cause.**

ENGELMANN-BECKER CORPORATION

A Quick Example: Students in Corrective Reading Making Short-Vowel, Long-Vowel Discrimination Errors
(Field Test 2023-2025 Maryland, Florida)

Students make vowel-discrimination errors by reading *slopping* as *sloping, hoped* as *hopped*, *stacker* as *staker, raked* as *racked*, *miles* as *mills*, and *thinner* as *thiner*.

This is precisely the kind of error pattern DI authors watch for. When an error pattern like this is detected, there are four main steps DI authors take to resolve the performance problem:

1. The Identification of the Domain and Range of the Error Pattern in the Field-Test Data
2. A Causal Interpretation of that Data
3. Author's Formulation and Implementation of a Solution
4. Iterative Refinements of the Proposed Solution Based on Field-Test Data from a new group using the updated version of the CR program.

After following the first two main steps, we realized that the error pattern was both pervasive and chronic, so we needed to introduce a reliable strategy.

In step 3, we decided to introduce the **vowel-sound rule**, which would provide students with a tool to make these vowel discriminations.

The first iteration of the of the **vowel-sound rule** mitigated, to a great extent, the error pattern but, in step 4, we further realized that the introduction of the rule had created a new error pattern as a side effect of curing the problem. When a new a group went through the revised version with the **vowel-sound rule,** students made errors applying the rule because specific pre skills had not been taught and firmed – identifying consonants between vowels.

The next iteration included more work on these pre skills before the introduction of the **vowel-sound rule.**

The final field-test groups went through this revised version with no consistent or noticeable performance side effects. Student performance on vowel-sound discriminations and all of the pre skills satisfied DI performance criteria.

Why DI Field-Testing Matters

Most programs claim to be "based on" or "aligned with" the Science of Reading. Direct Instruction is different:

- DI is the only instructional system that uses comprehensive, and complete student performance data to validate every component of instruction.
- Field-testing is not marketing.
- It is the iterative, technical process by which DI programs are built, refined, and verified.
- DI programs work reliably across student populations, instructional settings, and decades of research.

This process is what truly sets DI apart from everything else.