



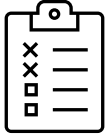
From Training Metrics to Service Quality:

A Data-Driven Approach to Improving Behavioral Support in I/DD Services

Jason St. Claire, MEd, MA
Emily Branscum, PhD, BCBA-D
Lee Ferguson, MSW, LICSW

The Challenge We Inherited

The context driving the project.



Concerns around the quality of positive behavior support (PBS) across various settings.

- Re: oversight entities, disability rights advocates, case manager observations, etc.



Providers often cite “*don’t know how*” as a central reason for subpar support.

- Providers indicate clinical plan writers are often confused by regulatory requirements and expectations related to evidence based best practices – and that this is the reason for the gaps in person-centered behavioral support.



Organizational pressures and systemic barriers create multi-layered challenges.

- Our directive: Do more PBSP writing trainings.
- Our starting point:
 - Substantial variation in practitioner experience, competency, and willingness to engage.
 - Vast logistical uncertainties + severely under-resourced regional clinical team infrastructure(s).
 - No vision for scope, metrics, evaluating progress / gauging impact, etc.

?!?!?!?



Recent policy revisions exacerbated provider uncertainties.

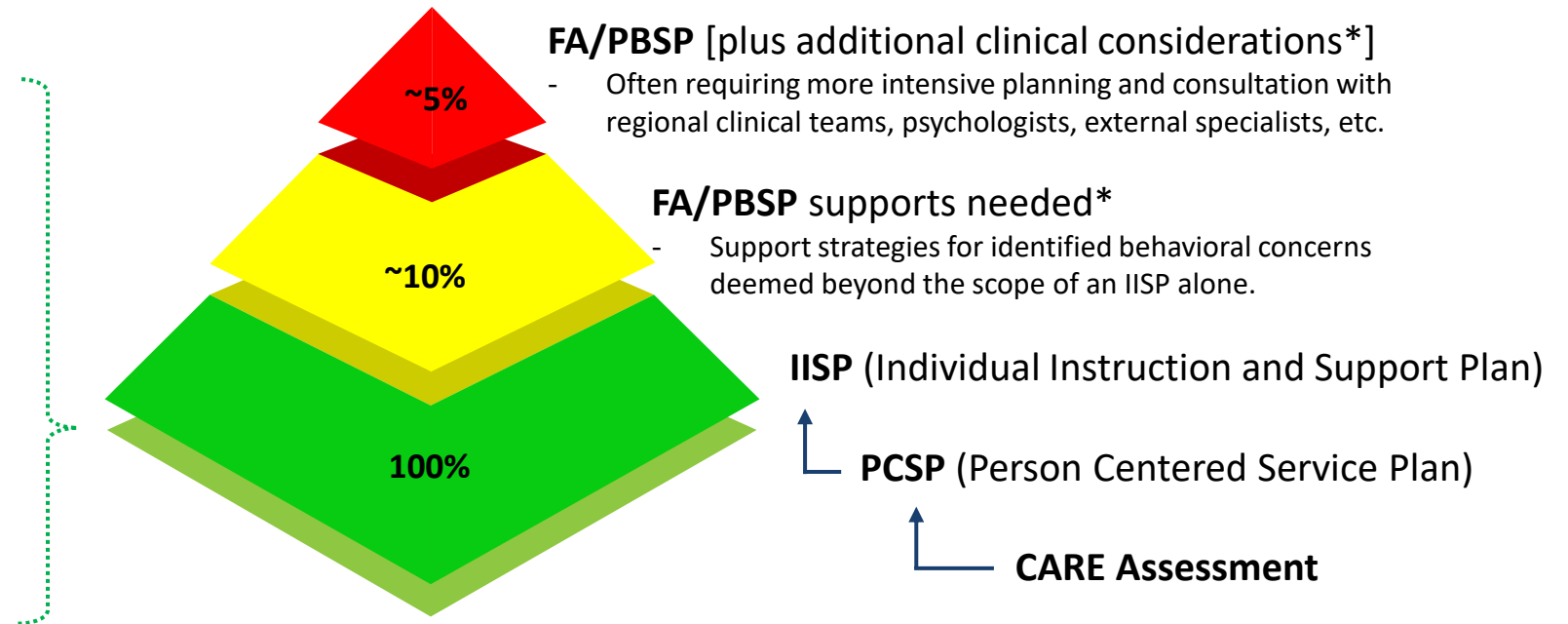
- Unclear follow up information intensified plan writer confusion – fomenting panic, urgency, and a crisis of confidence.

Determining Behavioral Support Needs

A quick note on scoping behavioral support needs and the tools we use.

DDA's guiding values:

- Inclusion.
- Status and contribution.
- Relationships.
- Power and choice.
- Health and safety.
- Competence.



Adding to the Challenge

Foundational questions we could not answer at the outset.

Will writing training work?

How will we know? How do we define and measure “*success*” beyond mere attendance & satisfaction survey data?

For whom?

Can we demonstrate equity of impact across the massive experience and educational variance in our practitioner population?

Where specifically?

Which phases or elements of plan development and implementation are breaking down in practice? Do we have clarity on this?

Based on what data or info?

Without baseline capability data, how can we know if anyone improved or moved in the right direction? What assumptions do we make?

Is there coherence across our systems?

Do the details of our training curriculum align with how we assess our training effectiveness?

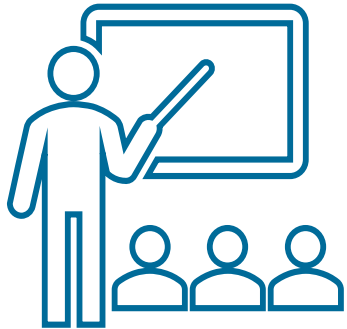
Does our approach to assessing training effectiveness align with what we review and evaluate for planning and implementation quality?



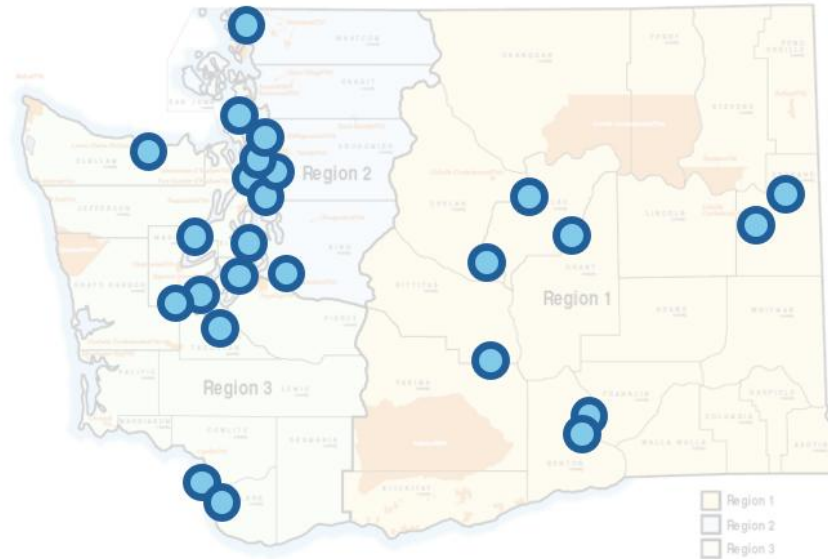
Without a clear signal between the specifics of what we train, how we assess training impact, and what we review & evaluate for quality – aren't we just assuming that mere attendance equals training “success”?

Our Response

20+ intensive, two-day trainings delivered across all three regions.



20+
Trainings



257+
Attendees

Systematic recalibration requires reaching practitioners where they are.



What's in the two-day training?

Day 1 – Functional Assessment

- Clinical reasoning fundamentals.
 - Hypothesis generation.
 - Identifying setting events & antecedents.
 - Data collection, analysis, and interpretation.
 - Adaptive skills analysis; etc.
- Antecedent-behavior-consequence mapping in cases.
- Policy requirements & clinical judgement (navigating tensions).
- Small-group breakouts, cases with facilitated problem-solving.
- Correcting field-wide misconceptions and/or ‘drift’ from best practices.
- Introduction to new format and template (how to utilize and navigate).

Day 2 – Positive Behavior Support Plan

- FA-to-plan integration (turning analysis into actionable support).
- Replacement behavior identification (functionally relevant).
- Teaching strategy development and approaches to reinforcement.
- Data collection approaches that work in real support environments.
- Person-centered language and implementation guidance for staff.
- Client rights. Informed consent. Restrictions / setting modifications.
- Template navigation, review schedule best practices, staff training.
- Resources, tools, materials and consultation pathways.

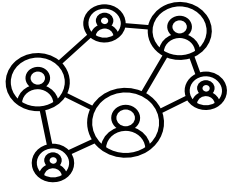
Each training needed to simultaneously address foundational knowledge, procedural application, policy compliance, and pervasive misconceptions – while introducing new information.



System-level complications presented a uniquely multi-dimensional challenge.

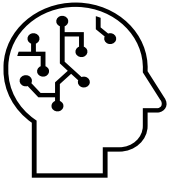
Designing a Curriculum for Everyone – From Novices to Veterans

Building a learning architecture that works for seasoned experts and new providers simultaneously – without losing either group.



Attendee variation as a significant design constraint.

- Experienced BCBA's sitting next to companion home providers with zero behavior support background.
- Can't pitch to the middle – both groups disengage.
- Need layered learning architecture; multiple entry points to the information



Multiple simultaneous learning objectives.

- Propositional knowledge (*what* is PBS? An FA? A replacement behavior? Data? Policy?).
- Procedural knowledge (*how* does one do or enact this? What is the process for 'x'?).
- Conditional application (*when/why* does this approach fit – practically speaking?).
- Template/format introduction and navigation (*what* it entails – how to utilize it; policy musts; etc.).



Policy tensions, resistance to change, and a heavily contaminated information ecosystem.

- Recent policy revisions heightened existing confusions, leading to proliferation of bad information.
- We needed to teach to the updates and confront widespread, deeply entrenched misconceptions.
- This required cleaning up the broader information landscape – while introducing new knowledge.

Facilitation as Real-Time Adaptive Problem Solving

Strategic approaches to keeping 250+ diverse learners engaged and calibrated.

The relevance problem.

- Making content feel essential, whether it is brand new to them or total review.
- Novices require foundations. Veterans needed both recalibration and fresh ideas.

Teaching to a broad variety of adult learners requires multiple modalities - simultaneously.

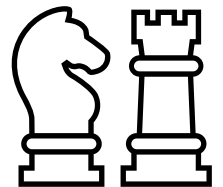
- Metaphor and analogy for conceptual learners – addressing language & educational gaps.
- Data and statistics for evidence-driven learners – overcoming resistances.
- Scenario walkthroughs and case examples – respecting practicality and the need for utility.

Navigating resistance without sacrificing rapport.

- Updated policies can cause anxiety around abandoning old practices and potential new requirements.
- Widespread and deeply-rooted misconceptions had to be tactfully confronted.

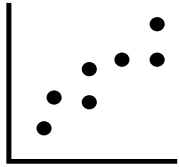
Cultivating community of practice.

- Providers often felt isolated; this training created a shared language and conceptual space.
- Deliberately emphasizing our shared community of practice, we unlocked vast amounts of distributed intelligence within the classroom.



The Assessment Framework: Professional Self-Efficacy

Elegance over extravagance. Signal over noise.



Our instrument – strategic, simplified, psychometrically sound.

- Maps directly onto the core clinical reasoning domains we teach and need to see in planning.
- 10-items, 1-5 scale, anonymous completion. No barrier to participation, no duress.
- Not a competency exam – a targeted, domain-specific professional self-efficacy diagnostic.

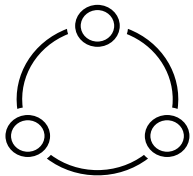


Tracks the appropriate thing - professional self-efficacy.

- We're measuring confidence to understand concepts and apply knowledge in real-world contexts.
- Professional self-efficacy routinely predicts transfer to practice better than knowledge tests alone.
- In this field, we are measuring *readiness*, not memorization.

Immediate instructive value – not just evaluative.

- Results show exactly where confidence gaps exist (pointing us toward our next teaching emphases).
- No waiting for external analysis. Data guides curriculum and delivery adjustments along the way.
- An active feedback loop; not a once-over. Dynamic iterative refinement at all levels.



Fully compatible with existing quality review approaches.

- Clinical reasoning domains align directly with quality review and evaluation criteria.
- Clear signal across our training curriculum, training assessment instrument, and quality review systems.

The Instrument: Essential Clinical Reasoning Domains

The connective tissue between what we teach, how we measure our effectiveness, and what we evaluate for quality.

- Conducting a functional behavioral assessment
- Collecting data related to target behaviors
- Identifying antecedents / 'triggers' for behavior
- Collecting data related to replacement behaviors / skills
- Analyzing behavioral data (toward refining the plan)
- Writing in person-centered language/ avoiding pejorative language
- Turning the functional assessment into a practical usable PBSP
- Developing teaching/training strategies
- Identifying replacement behaviors
- Knowing who to contact for additional consult regarding complex cases

	1	2	3	4	5	
	Not at all confident	Slightly confident	Somewhat confident	Quite confident	Extremely confident	
Conducting a functional behavioral assessment.						Pre-
Collecting behavioral data for target/challenging behaviors.						
Identifying setting events / antecedents (triggers) for behavior.						4 5 Post-
Collecting behavioral data for replacement behaviors / skills development.						4 5
Analyzing behavioral data (toward refining the plan, based on data).						4 5
Writing in person-centered language / avoiding pejorative language.						4 5
Turning a functional assessment into a practical, usable PBSP.						4 5
Developing teaching/training strategies (toward development of new skills).						4 5
Identifying replacement behaviors.						4 5
Knowing who you can contact for guidance/consultation around complex cases.						4 5
Knowing who you can contact for guidance/consultation around complex cases.						4 5

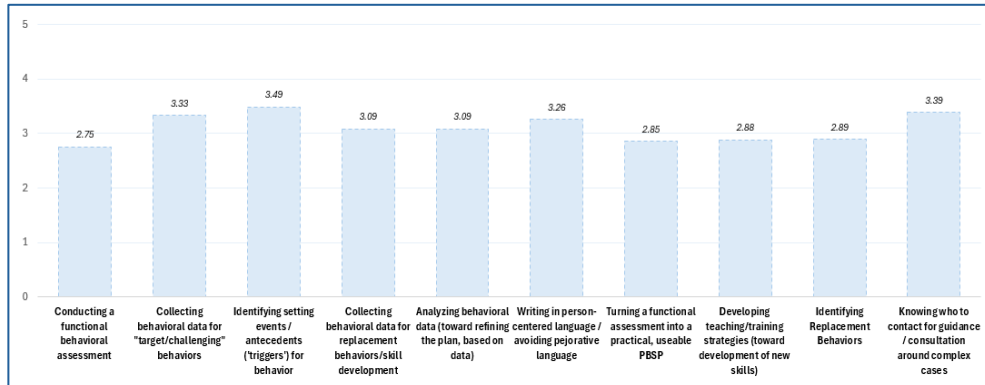
Total Score = "Confidence Quotient" (CQ): ___ / 50

This approach provided us with...

Direct Comparison ‘Interventional’ Data

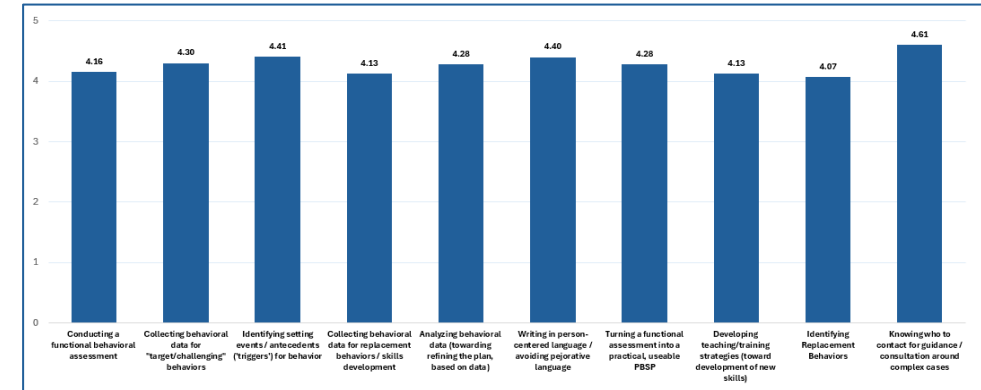
Professional self-efficacy – by and across clinical reasoning domains – for each class.

Pre-



Domain-specific professional self-efficacy walking in...

Post-



Domain-specific professional self-efficacy walking out...

- This was not us conducting an evaluation from the outside-looking-in. No projection.
- We invited attendees to tell us about themselves – for *specific* areas of professional competence – both before and after our ‘intervention.’
- Built on a simple formula: deep, incisive inquiry + sincere listening + dynamic refinement via precise data collection and analysis.

Giving Rise to 'Reports' Following Each Training

Initiating a high-fidelity, data-driven feedback loop to inform and sharpen our efforts along the way.



Over the course of more than a year –
across **20(+)** training cycles and **257(+)**
practitioners – clear patterns emerged.

An aerial photograph of a dense evergreen forest, likely a spruce or fir forest. The trees are packed closely together, creating a textured, dark green canopy. The lighting is bright, highlighting the individual tree tops and creating a sense of depth. The text "First, a quick flyover..." is centered in the middle of the image in a white, sans-serif font.

First, a quick flyover...

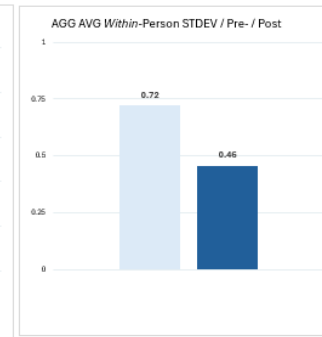
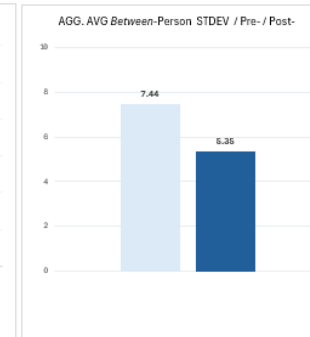
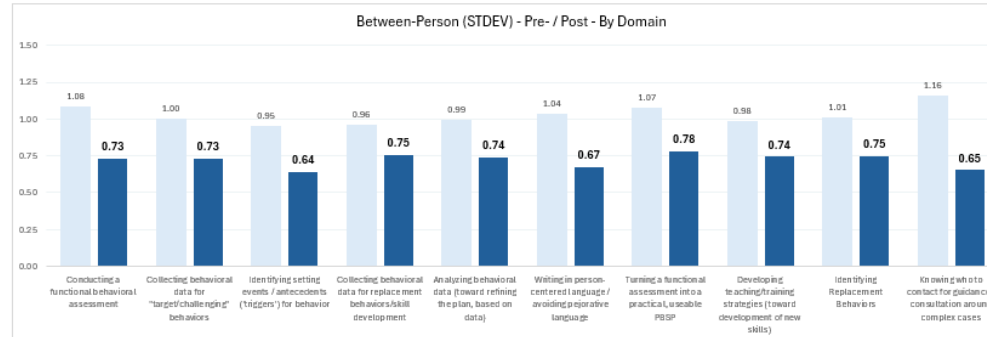
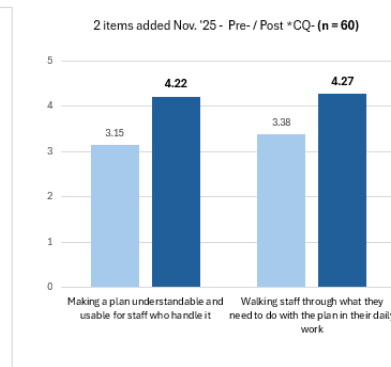
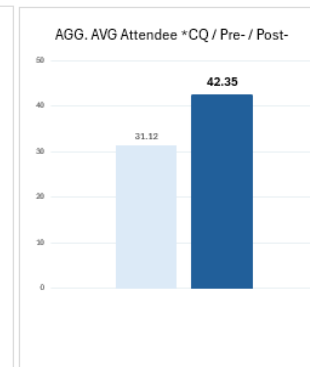
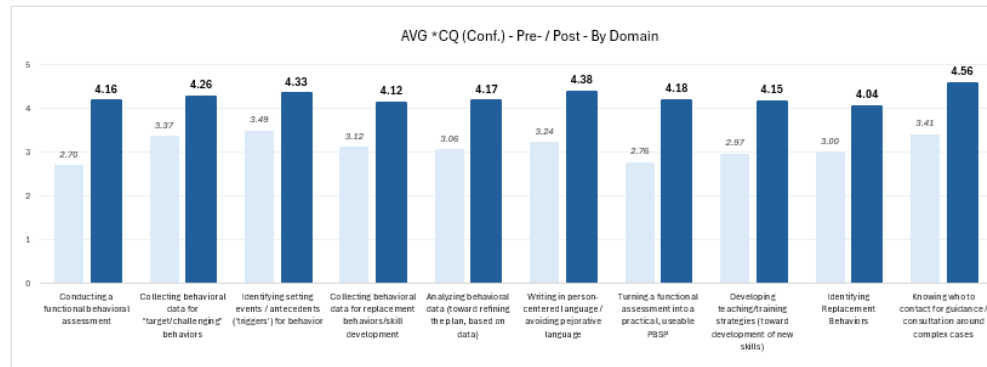
The 10,000 ft. View: Noticing the Forest Before the Trees

Confidence surged across all clinical reasoning domains. We will examine details momentarily. First, notice the pattern.



n = 257

Light = Pre- Dark = Post-

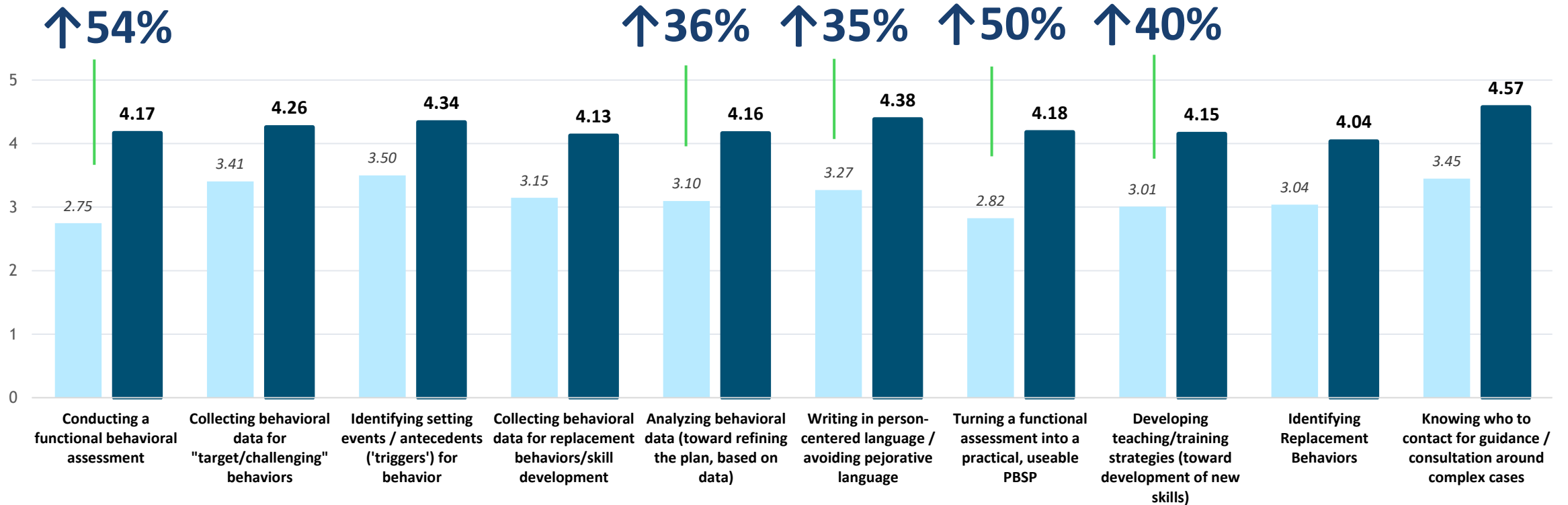


A wooden bridge with railings spans across a stream in a lush, green forest. The bridge is made of weathered wooden planks and has simple wooden railings on both sides. The surrounding area is filled with dense foliage, including various green leaves and ferns. The stream is visible beneath the bridge, with some rocks and water. The overall scene is peaceful and natural.

Now, let's zoom in...

Every Domain Saw Significant Growth – Some Doubled.

Across all domains the average practitioner confidence increase exceeded 35%.



AVG Conf. - Pre- / Post - By Domain

n = 257



A different look...

Double-Digit Growth Across All Domains

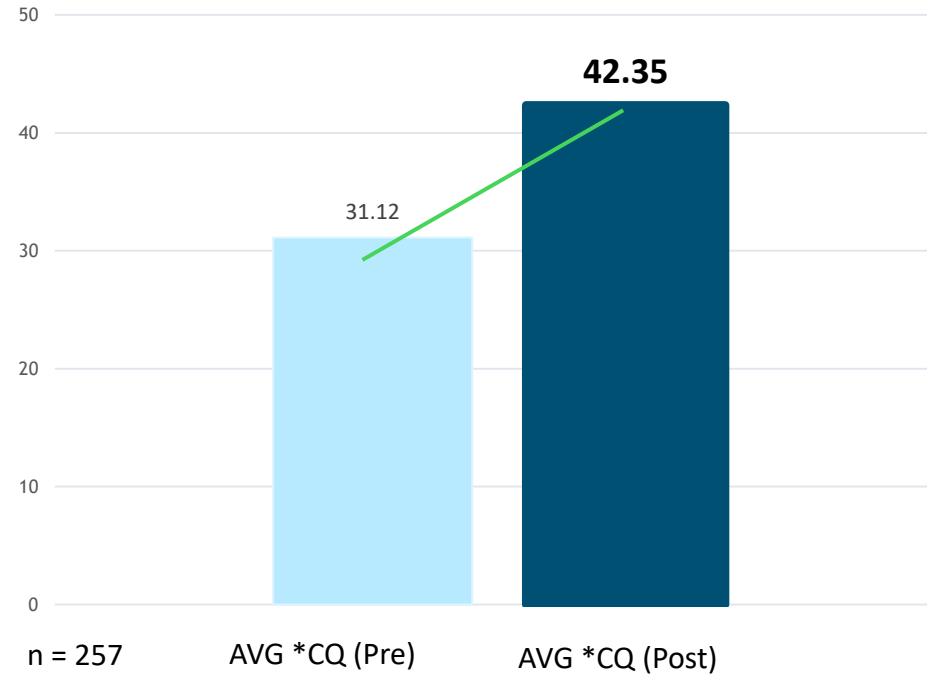
Clarifying how it all fits together – and seeing it show up decisively in the data.

Clinical Reasoning Domain	Confidence / Self-Efficacy		
	Avg. CQ		
	Pre-	Post-	Shift
Conducting a functional behavioral assessment	2.71	4.17	↑ 54%
Collecting data related to target behaviors	3.38	4.26	↑ 26%
Identifying antecedents / 'triggers' for behavior	3.49	4.34	↑ 24%
Collecting data related to replacement behaviors / skills	3.12	4.13	↑ 32%
Analyzing behavioral data (toward refining the plan)	3.07	4.17	↑ 36%
Writing person-centered / avoiding pejorative language	3.24	4.39	↑ 35%
Turning FA into practical usable PBSP	2.79	4.18	↑ 50%
Developing teaching/training strategies	2.97	4.16	↑ 40%
Identifying replacement behaviors	3.00	4.04	↑ 35%
Knowing who to contact for additional consult regarding complex cases	3.44	4.58	↑ 33%

n = 257

Overall Practitioner Confidence Rose Dramatically

Moving the average attendee from the 50th to 80th percentile.



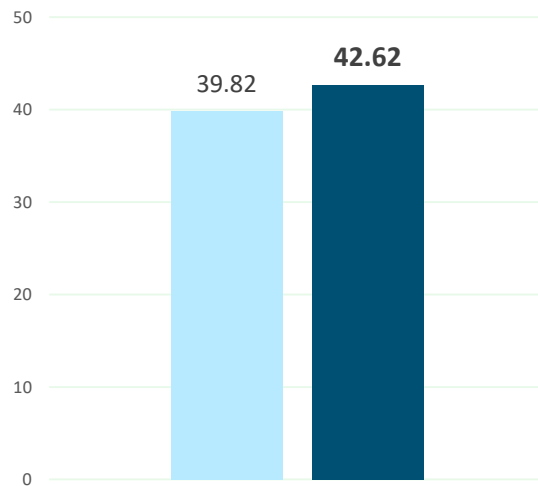
↑ 36%

From mid-range to top quartile level.

Equity and Differential Impact: Closing the Confidence Gap

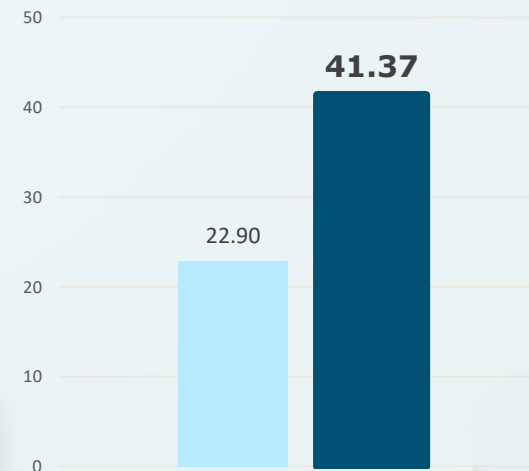
Confidence is up across the board - with the most pronounced gains showing up where they were needed most.

↑ 7%



High Conf. (defined as total pre- CQ score >34) [n = 78]
7% Confidence boost (*ceiling effect)
13% of high conf. attendees reached post-CQ score of 50.

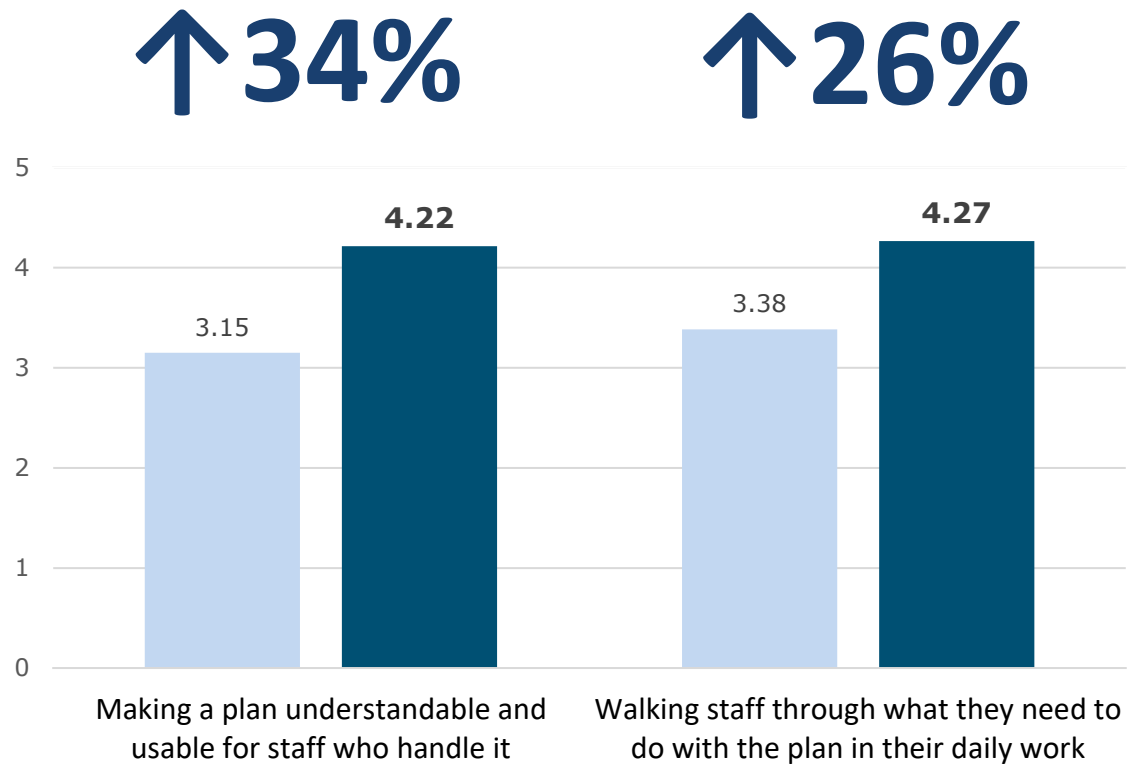
↑ 81%



Low conf. (defined as total pre- CQ score ≤28) [n = 86]
81% Confidence boost (for most in need)
73% of low conf. attendees reached post- CQ score of 40+.

Real-Time Refinement: System Response to Field Data

Two new domains were added mid-project (Nov. '25) based on patterns in feedback; sizable gains quickly materialized (n = 60).



...but increased practitioner confidence is the *least* interesting part of the findings...



It is reasonable to ask:

Is this just a handful of tall trees – or forest-wide growth?

What about variation between attendees? Between domains?

What about all the beginners / novices? Were they left behind?



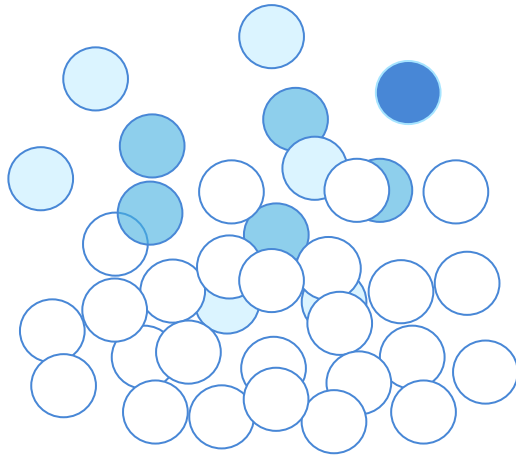
Let's take a closer look...

...after a quick refresher...

What is 'variance' or 'variation' in this context?

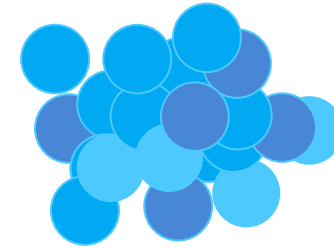
High variance

- Up in some areas, down in others
- *All over the map / mixed-bag / spread out*
- Big differences between respondents



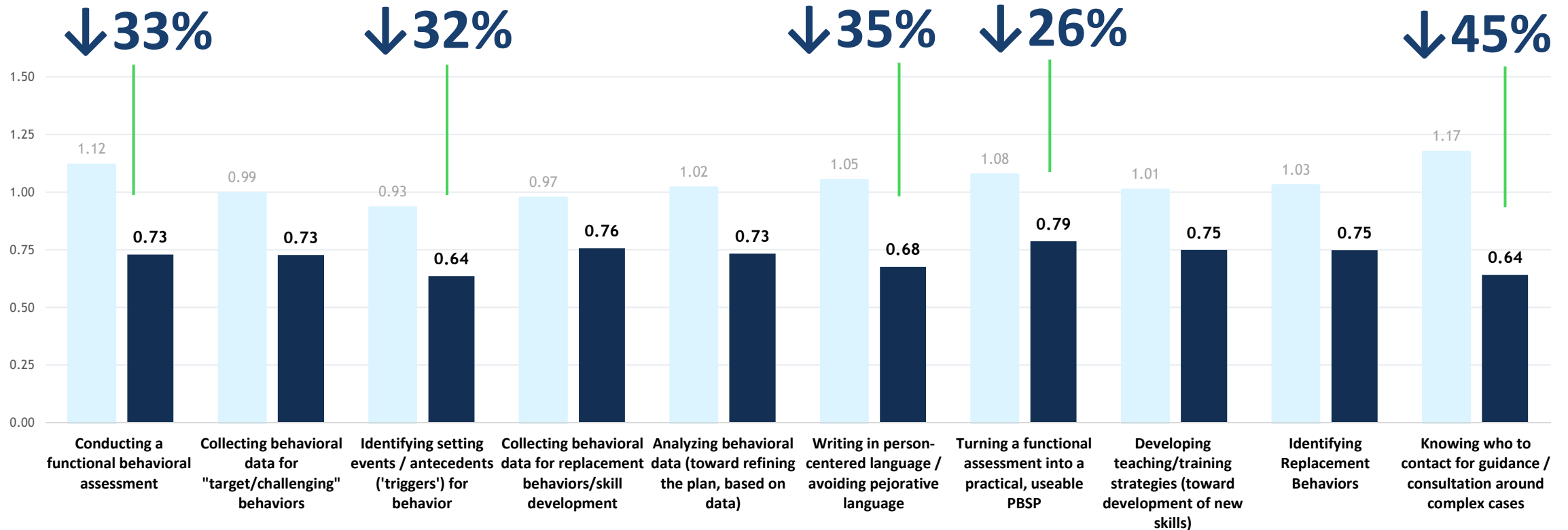
Low variance

- Convergent. Cohesive. Tightly aligned.
- More even / balanced / rounded across domains.
- Little difference between respondents



Variance Collapsed: Up to 45% Boost in Practitioner Convergence

Across all domains, scattered confidence became considerably more aligned (avg. 29%).



Between-Person (STDEV) - Pre- / Post - By Domain

n = 257

A Different Look: Narrowing the Competency Divide

Practitioner convergence boosted by an average of **29%**. Like raising the ceiling – and the floor as well.

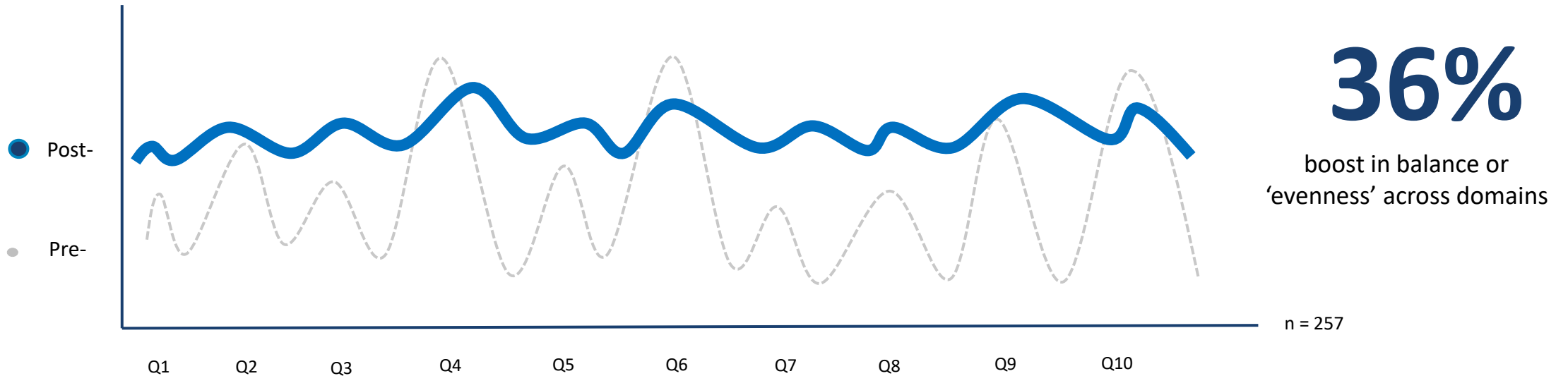
Clinical Reasoning Domain	Between-Person Variance		
	AVG.		
	Pre-	Post-	Shift
Conducting a functional behavioral assessment	1.09	.73	↓33%
Collecting data related to target behaviors	.98	.73	↓26%
Identifying antecedents / ‘triggers’ for behavior	.93	.64	↓32%
Collecting data related to replacement behaviors / skills	.96	.76	↓21%
Analyzing behavioral data (toward refining the plan)	1.00	.73	↓27%
Writing person-centered / avoiding pejorative language	1.04	.68	↓35%
Turning FA into practical usable PBSP	1.07	.79	↓26%
Developing teaching/training strategies	.99	.75	↓24%
Identifying replacement behaviors	1.01	.75	↓26%
Knowing who to contact for additional consult regarding complex cases	1.16	.64	↓45%

n = 257

Compressing the distribution – strengthening collective capability.

Balanced Self-Efficacy – Not Just Isolated Areas

Decisive improvements in “evenness” of confidence *across* clinical reasoning domains.



- ❖ **Pre-training**, individuals had imbalanced confidence profiles – strong (*high*) in some areas, vulnerable (*low*) in others. Variation.
- ❖ **Post-training**, confidence profiles showed consistent elevation across *all* domains – **reducing imbalances by 36%**.
- ❖ **Translation**: substantially reduced spread = well-rounded, more balanced, steady confidence *across* skill domains.

...but this is *still* not the most interesting part of the findings...



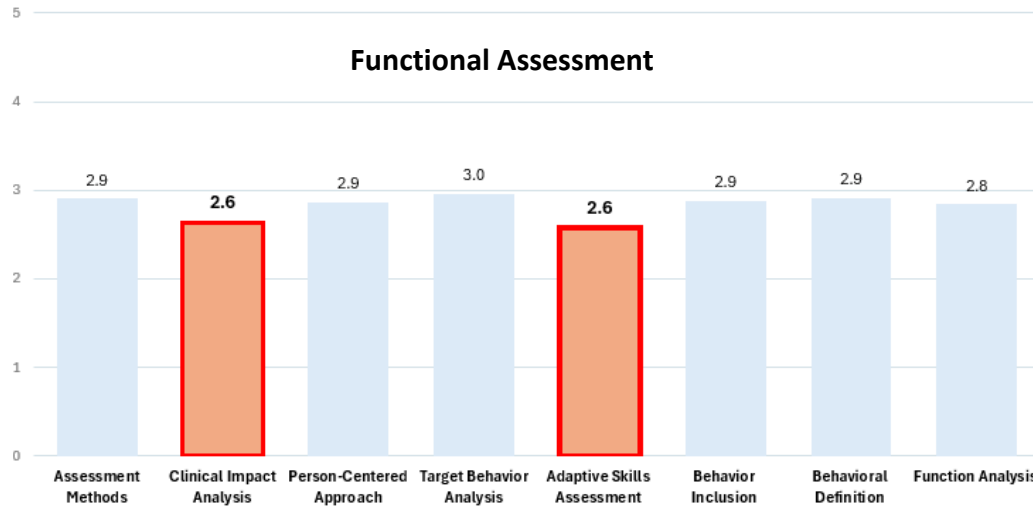
...throughout 2025, we had been experimenting with a system for enhanced review analytics – compiling a diagnostic database of where plan development struggles most...



...certain domains kept appearing as
the most persistent vulnerabilities.

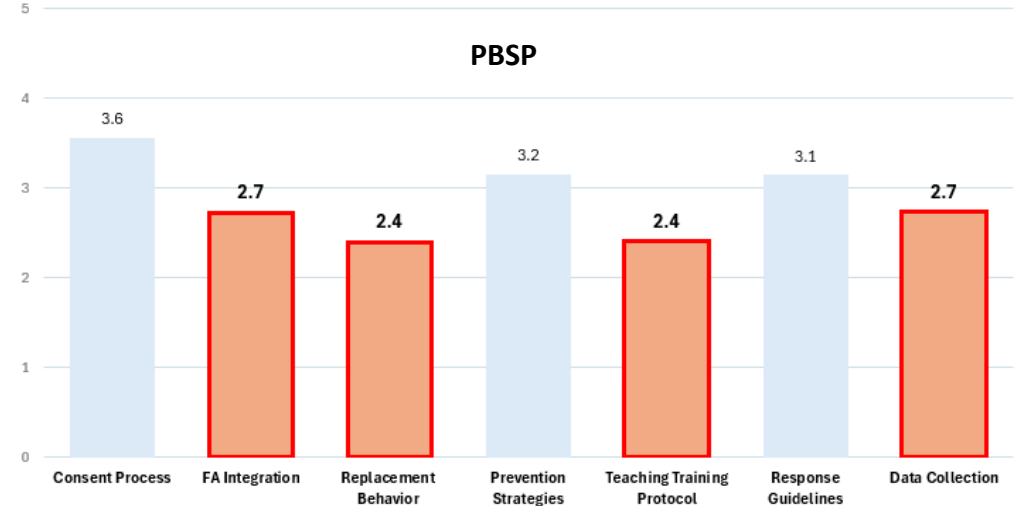
'Pain Points' – Mapping Common Pitfalls

Clinical quality review data revealed persistent areas of struggle. These are not random gaps.



Clinical Impact Analysis

Adaptive Skills Assessment



FA Integration (into PBSP)

Replacement Behaviors

Teaching/Training Strategies

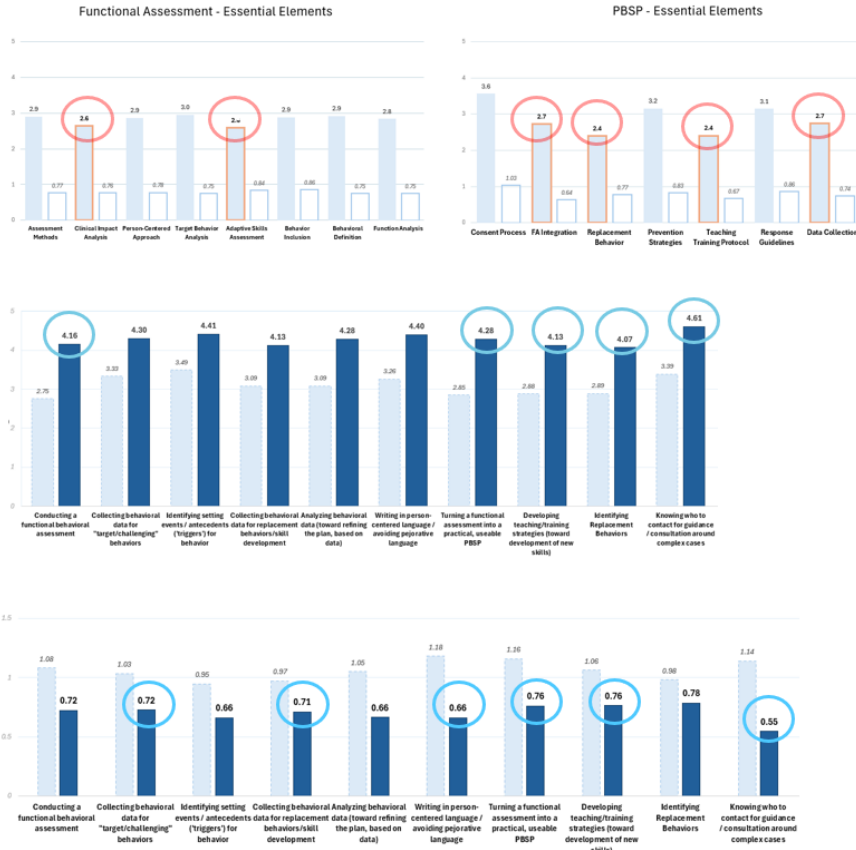
Data Collection (General)

- These 'pain points' map directly to detailed, domain-specific review criteria.
- This enabled us to measure patterns and trends that had otherwise only been alluded to.
- These are not trivial gaps; they represent foundational vulnerabilities that demand targeted attention.

How did our training project results measure up against these vulnerabilities?

Direct Hits.

Several key 'pain points' ended up as the areas of greatest improvement.



'Pain Points' – Top 5 Identified:

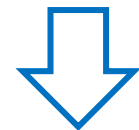
1. Clinical impact / adaptive skills (conducting FA)
2. Turning FA into useable PBSP
3. Teaching/Training Strategies
4. Replacement behaviors
5. Data Collection (generally)

Confidence Growth – Top 5 Domains:

1. Conducting FA
2. Turning FA into useable PBSP
3. Teaching /Training Strategies
4. Identifying replacement behaviors
5. Knowing who to contact for guidance

Variance Reduction – Top 5 Domains:

1. Turning FA into useable PBSP
2. Teaching/Training Strategies
3. Data Collection
4. Writing in person-centered language
5. Knowing who to contact for guidance



C
O
N
G
R
U
E
N
C
E

This congruence demonstrates precision
and is highly encouraging.

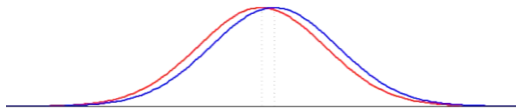
...but what about actual effect size?

Did we make it to the roots?

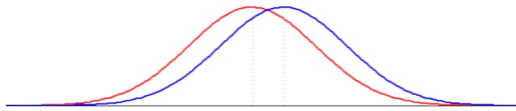
First...a quick refresher...

Standardized Effect Size (*Cohen's d*)

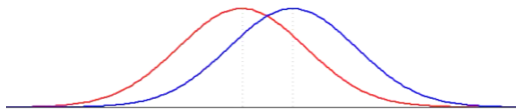
- Simple metric for determining whether a change is small, moderate, or big enough to matter.
- Measures how far apart two group averages are using a common statistical yardstick. In our case, between *pre-* and *post-*.
- Translates the difference between pre- and post- scores into a standardized scale that shows us *magnitude* of change.



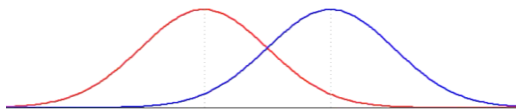
$d = 0.2$: Small effect (subtle movement)



$d = 0.5$: Medium effect (noticeable movement)



$d = 0.8$: Large effect (substantial movement)



$d = 2.00$ Enormous effect (exceptional movement)



Now, let's take another look at our results...

The Full Signal.

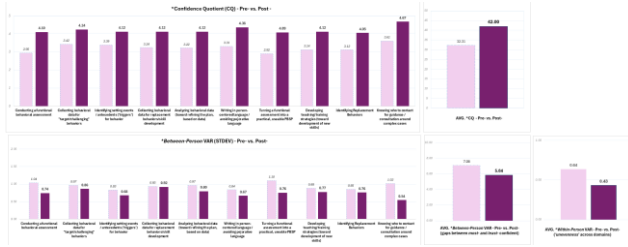
All domains saw considerable confidence growth and practitioner alignment – stamped by robust effect sizes.

Clinical Reasoning Domain	CQ	STDEV	<i>d</i>
	<i>Professional self-efficacy ↑</i>	<i>Between-person variation ↓</i>	<i>Standardized effect size</i>
Conducting a functional behavioral assessment	↑54%	↓33%	*1.58
Collecting data related to target behaviors	↑26%	↓26%	*1.02
Identifying antecedents / ‘triggers’ for behavior	↑24%	↓32%	*1.07
Collecting data related to replacement behaviors / skills	↑32%	↓21%	*1.16
Analyzing behavioral data (toward refining the plan)	↑36%	↓27%	*1.25
Writing person-centered / avoiding pejorative language	↑35%	↓35%	*1.31
Turning FA into practical usable PBSP	↑50%	↓26%	*1.49
Developing teaching/training strategies	↑40%	↓24%	*1.35
Identifying replacement behaviors	↑35%	↓26%	*1.17
Knowing who to contact for additional consult regarding complex cases	↑33%	↓45%	*1.22

* >.80 = Large effect size.

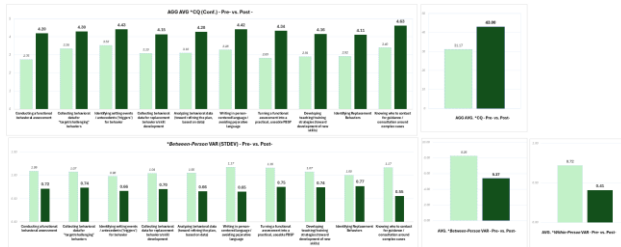
Regional Comparisons - A Nod to Consistency

R1



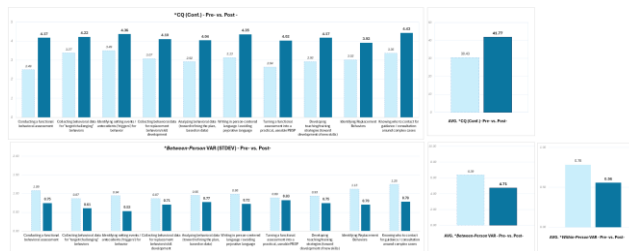
Confidence: 30.4% average practitioner confidence increase
Alignment: 16.9 % reduction in readiness gaps between high and low practitioners. Equity.
Balance: 33.3 % boost in practitioner evenness *across* domains. Balanced self-efficacy.
Effect size (range): *.91 – 1.41(!) – Strong effects across all domains.
Reliability: 86% - [14% CV] – High consistency across individual trainings.

R2



Confidence: 37.9% average practitioner confidence increase
Alignment: 34.9 % reduction in readiness gaps between high and low practitioners. Equity.
Balance: 43.3 % boost in practitioner evenness *across* domains. Balanced self-efficacy.
Effect size (range): *1.02 – 1.57(!) – Strong effects across all domains.
Reliability: 87.5% - [12.5% CV] – High consistency across individual trainings.

R3



Confidence: 37.3% average practitioner confidence increase
Alignment: 26.2 % reduction in readiness gaps between high and low practitioners. Equity.
Balance: 29% boost in practitioner evenness *across* domains. Balanced self-efficacy.
Effect size (range): *.95 – 1.80 (!) – Strong effects across all domains.
Reliability: 88.7% - [11.3% CV] – High consistency across individual trainings.

A dense forest of evergreen trees, likely spruce or fir, is shown under a heavy rain shower. The rain is falling diagonally across the frame, creating a sense of movement and atmosphere. The trees are dark green and densely packed, filling the entire background. The overall mood is serene and natural.

Not just improvements – recalibration at scale.

Deep saturation in a field desperate for grounding.



This is no coincidence.

We detected core vulnerabilities or ‘pain points’

(via intensive review, data collection, and analysis)

We deliberately targeted these areas with focused emphasis throughout the trainings' core curriculum
(e.g., additional tools / materials / resources / activities)



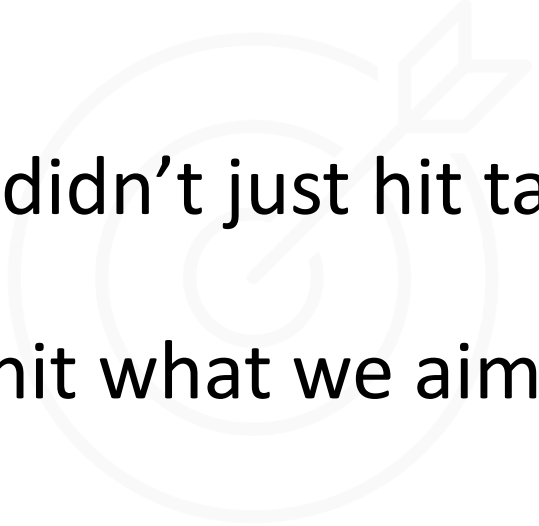
Facilitated, small-group
breakout activities



- Sample/mock plans
- Concrete examples

Conceptual models / clarification of frameworks / guiding activities / practical tools, etc.

- Data analysis examples



We didn't just hit targets.
We hit what we aimed for.

Our framework: an integrated systems approach to training data – from development and delivery through analysis and reporting, and back again.

Underlying Philosophy: Dual-Focus Design

One integrated conceptual framework addressing both practitioner capability *and* implementation context

Plan Development Capability

(what's needed to develop higher quality behavior support plans)

Our emphases:

- Precision. Domain clarity.
- Foundations. First principles.
- Shared conceptual language.



Plan Implementation Fidelity

(what's needed to increase the likelihood of *delivery* of quality support)

Our emphases:

- Advocacy
- Client rights / informed consent
- Universal supports

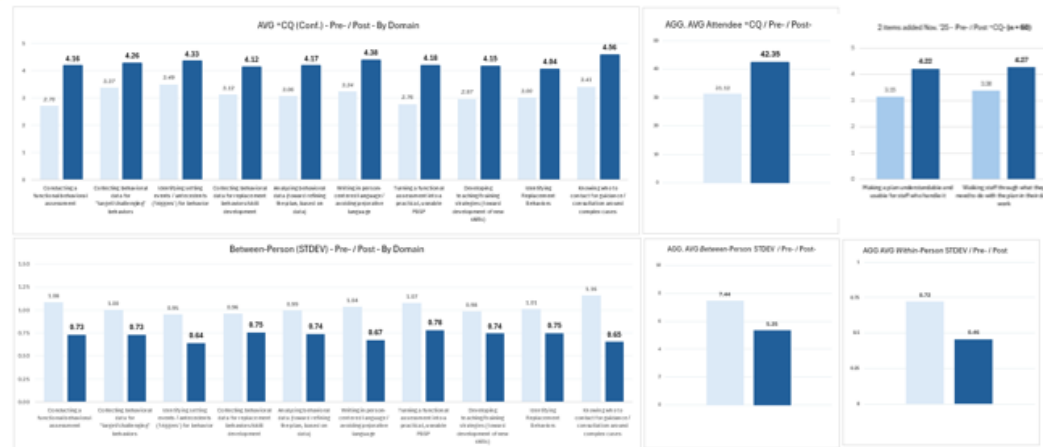
We didn't inherit an operable, instructive, or generative system to help us in confronting this sprawling, multi-dimensional project.

So, we engineered one.

From vision to viability.



Back to the big picture of the day – hard data.



Remember Those Foundational Questions from Earlier?

Now we can confront them.

- Will writing training work?**
How will we know? How do we define and measure “*success*” beyond mere attendance & satisfaction survey data?
- For whom?**
Can we demonstrate equity of impact across the *massive* experience and educational variance in our practitioner population?
- Where specifically?**
Which phases or elements of plan development and implementation are breaking down in practice? Do we have clarity on this?
- Based on what data or info?**
Without baseline capability data, how can we know if anyone improved- or moved in the right direction? What assumptions do we make?
- Is there coherence across our systems?**
Do the details of our training curriculum align with how we assess training effectiveness?
Does our approach to assessing training effectiveness align with how we review and evaluate planning and implementation quality?



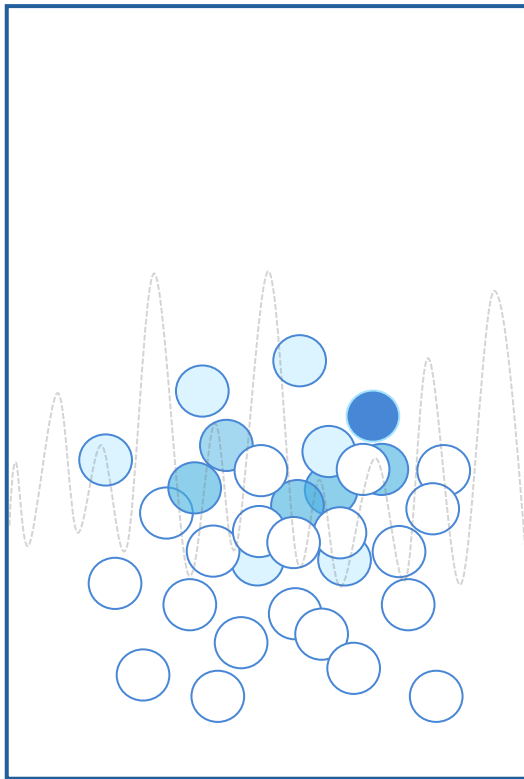
**Do we have a framework within which our trainings
can be systematically assessed as impactful?**

The Signal Emerging From the Noise:

Key insights and remaining challenges

Where we were

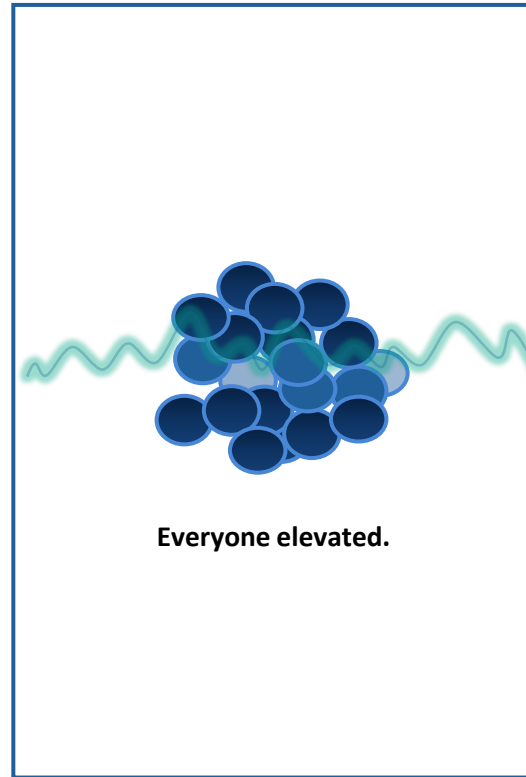
(High variance, low-to-mid confidence)



Some got it – many were tentative.

Where we are

(low variance, high confidence)



Everyone elevated.

The entire cohort signals capability.

What's next?

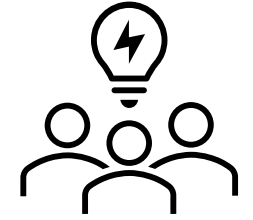
The implementation challenge



Knowledge gaps closed – other barriers now visible.

What the data tells us – and what it doesn't.

Four things that matter.



1. The knowledge gap closed – considerably.

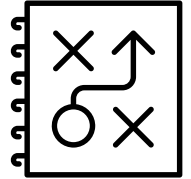
The finding:

- **257(+)** practitioners via **20(+)** trainings across all 3 regions
- **36%** average practitioner confidence increase. Multi-domain professional self-efficacy.
- **29%** reduction in variability between practitioners (bringing *everybody* up)
- **36%** reduction in practitioner variability across clinical domains (boosting balance of skill profile)
- Strong effect sizes across all domains (*d-range = 1.02 – 1.58*)

What this means:

The “don’t know how” narrative for PBS quality gaps is no longer defensible.

Field wide recalibration was achieved by way of a simple formula: unified messaging from clinical subject matter experts, emphasis on fundamentals of best practice, and strategic correction of widespread misconceptions (i.e., cleaning up the information landscape). This was broad spectrum demystification.



2. We hit exactly what we targeted – systematically.

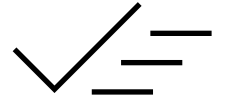
The finding:

- Quality review data identified at least six persistent pain points across a sample of 50 in-depth plan reviews.
- Training gains aligned precisely with these vulnerabilities; the top 5 areas of improvement mapped directly to core identified weaknesses.

What this means:

When curriculum, assessment, and evaluation share a unified conceptual framework, training precision can increase dramatically.

This wasn't another generic workshop series – it was targeted intervention informed by diagnostic professional self-efficacy data.



3. The framework proved itself – ecologically valid and verifiable.

The finding:

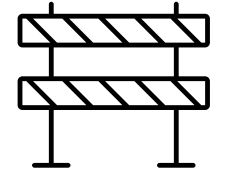
- We developed an integrated measurement ecosystem where professional self-assessment, training curriculum, and quality review share conceptual DNA and an emphasis on domain specificity. This gave us precision and clarity around where we needed to focus most and to what extent or depth.
- This kickstarted a feedback loop of educational coherence, richly instructive data, and actionable signal.
- This approach is generalizable, requires no specialized software, and could be replicated by anyone with clarity of vision and sufficient analytical rigor.

What this means:

Our approach shows what's possible when training infrastructure is built from first principles – with systems thinking, diverse subject matter expertise, and commitment to precision driving both design and operation.

For this project, re-imagining meant returning to abandoned foundations and excavating the cornerstones.

They still work. We built on them intentionally.



4. Implementation barriers become increasingly clear.

The finding:

- With foundational knowledge established statewide, we now have unprecedented clarity about what enables – or constrains – effective positive behavior support in practice:
 - Fundamentals of the support milieu (staffing consistency, daily routines, team culture)
 - Provider onboarding and caregiver training systems
 - In-home leadership presence and oversight / implementation fidelity monitoring
 - Resource allocation patterns (time, materials, operational capacity)

What this means:

These aren't curriculum gaps – these are organizational and operational realities that knowledge transfer alone can't solve.

Persistent quality concerns point clearly to structural enablers: the conditions, systems, and resources that determine whether practitioner capability translates (or *can* translate) to the fundamentals of person-centered support.

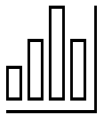
What This Means – Assets & Capabilities

What do we have now that we didn't at the outset?



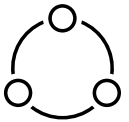
Measurable Return

- ✓ We now have demonstrable ROI data when providers ask, *“What is the state doing to help our behavior specialists?”*
- ✓ Receipts. *“Here’s how you sent your people to us – here’s how we sent them back.”* Measurable, trackable transformation.



Diagnostic Baseline Across the Field

- ✓ A comprehensive capability profile of 257+ practitioners across all regions. Granular data across all domains.
- ✓ Quality review can now better distinguish between knowledge gaps (rare) and implementation barriers (common).
- ✓ RCT consultation can now build on broadly established foundations rather than starting from zero each time.
- ✓ We now have a ‘pulse’. We see where the field stands – not just where individual practitioners tend to struggle.



Replicable Methodology

- ✓ A proven framework for data-informed training that can adapt, improve, scale, and generalize.
- ✓ Nested feedback loops enable continuous refinement based on real multi-source data.
- ✓ This is proof of concept for precision-targeted professional development in our field.

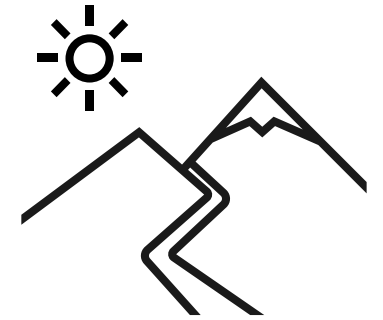


An Enhanced Standard of Accountability to the People We Serve

- ✓ The people we support deserve practitioners who are genuinely prepared. This system moves us measurably toward that standard.
- ✓ We can now trace from training room to residential setting – from what providers learn, to how support unfolds.
- ✓ Every confidence point gained in the training has a downstream address: someone’s home, someone’s day, someone’s life.
- ✓ The individuals we serve now have a broad constituency of better-prepared practitioners advocating for them.

Where We Go From Here...

- ❖ Naturally, our focus shifts more toward **implementation fidelity** – as our aim is to increase the consistency with which these supports reach people *in their homes*.
- ❖ This means **continued emphasis** on:
 - *Fundamentals of a successful living/learning environ. (healthy support milieu)*
 - *Trauma informed care*
 - *Client rights advocacy*
 - *Data collection / data driven support*
- ❖ We **continue to extend our reach** – working with providers on how they can better teach, monitor, and sustain individualized supports *within* their organizations.
- ❖ We **continue to coordinate** with community-based specialists and mental / behavioral health providers – reinforcing our commitment to person-centered whole health.
- ❖ On the momentum of what we’ve learned, we **continue to develop new training, workshop, and info-session opportunities**.
- ❖ Crucially, we didn’t “finish” something here. We established an approach worth replicating – and left the door open for anyone who comes next. This isn’t a conclusion. It’s infrastructure.



The background of the slide is a vibrant photograph of a coniferous forest. The trees are tall and slender, with dense green needles. The sky is a clear, bright blue, filled with soft, white cumulus clouds. The foreground shows a mix of green grass and small shrubs, suggesting a healthy, well-maintained forest environment.

Thank you.

Methodological Notes & Definitions

“Confidence Quotient (CQ)”: Simple proxy measure – Sum across all clinical-reasoning domains - each attendee. Total score = *CQ (___ / 50)

Between-person variance (STDEV): The space or gaps between attendees’ confidence scores (some people high, many people low).

Within-person variance (STDEV): The space or gaps *within* attendees’ confidence across the 10 clinical domains. Skill profile peaks and valleys.

Psychometric reliability rating: Measured by *Cronbach alpha* (.90 pre- / .91 post-). Translation: Excellent instrument reliability.

Standardized effect size (Cohen’s *d*): Simple measure to clarify magnitude or size of change.

Coefficient of Variance (CV): How reliable are the results? How replicable? We want this number low (= less variation between trainings).

Most crucially:

- Where’s the data coming from?
- This report is not us conducting an evaluation of practitioners from the outside-looking-in.
- This is what attendees are telling us about themselves – for *specific* areas of competence – both before and after our ‘intervention’.
- Born from a simple formula: deep, incisive inquiry + sincere listening + dynamic refinement via precise data collection & analysis.