THE PHASE 2 MANUFACTURING WALL™ · COMPANION INSTRUMENT

Facilitator's Guide to Honest Scoring

How to run the Readiness–Exposure Gap™ assessment so it tells you the truth.

Lonnie Robarge, Ph.D. · Bharat Gurale, Ph.D.

REVISED FOR THE 6×6 INSTRUMENT · REV. 2

01 — WHY THIS GUIDE EXISTS

The instrument's one failure mode is a flattering score.

The Readiness–Exposure Gap™ has one failure mode, and it is not a high score. It is a flattering one. A diagnostic scored by the same people whose work it measures will drift upward unless the scoring is deliberately governed — the Illusion of Readiness quietly reasserting itself inside the very instrument built to expose it. A program that scores itself a comfortable Intervention Window when it is actually At the Wall has not passed the assessment; it has defeated it — and it will meet the consequences later, when they are far more expensive.

This guide exists to make the inputs trustworthy. The arithmetic is simple; the honesty is hard. What follows is how to run the assessment so the result reflects what the organization actually knows, not what it hopes.

02 — THE GOVERNING PRINCIPLE

Visibility, not judgment.

Before a single dimension is scored, the room has to believe one thing: finding a gap is the goal, not a failure. Unresolved manufacturing exposure surfaced early is exactly what the instrument is for — it is risk that already exists, now made visible while it is still affordable to address.

> **If the people in the room suspect the score will be used to judge them, they will protect themselves by inflating it — and you will harvest a comfortable, useless number.**

This is not a motivational aside; it is mechanical. The sponsor must say, in plain words and at the very start: **we are looking for where we are exposed, not grading anyone, and the best outcome today is an honest gap we can act on.** Only then does the scoring have a chance of being true.

03 — BEFORE THE SESSION

Set the room before the first score.

Frame it out loud.
The sponsor or most senior person opens by stating that a high gap found now is a win, and that no score will be held against any function. Say it explicitly — silence is read as threat.

Build the right room.
Score with a cross-functional group, not leadership alone. Process development and MSAT must be present — they hold the tacit, undocumented knowledge the instrument is trying to externalize, and several dimensions cannot be scored honestly without them. Analytical, quality, clinical, and finance perspectives all belong in the conversation.

Appoint a neutral facilitator.
The facilitator should not be the program lead and should hold no stake in the score. Their only job is to ask, again and again, **what is the evidence for that?** A facilitator defending the program cannot govern the scoring.

04 — THE CARDINAL RULE

Evidence over confidence.

This is the heart of honest scoring, and most of the work. A score is a claim about what you know, not about how well things have gone. Repeated batch success is not evidence of robustness — it is the framework's central warning. A process can run flawlessly inside a narrow window for years while remaining completely uncharacterized outside it.

> **Apply one test to every score of 3 or 4: “What artifact would I put in front of a skeptical regulator, or an acquirer's diligence team, to defend this?”**
> **If the honest answer is “things have gone fine,” that is a level 1, not a level 3. No data, no high score.**

The artifact test governs the six Understanding dimensions. The Exposure axis is scored differently — by how much dependency has actually accumulated, not by evidence — so never demand an artifact for how much capital, timeline, or regulatory position is committed. On the Exposure side, a high score is a fact about the enterprise, not a claim to be defended.

> **Design-space characterization.** Level 3 means data across a design space — ranges, edges, interactions. Consistency at a validated setpoint, however reliable, is level 1–2. No DoE or range data caps the score at 2. *(This dimension now absorbs the former “operating-boundary knowledge”; behavior known across ranges is part of the same evidence.)*

> **Impurity & degradation knowledge.** Level 3 means formation pathways are understood mechanistically. Observing acceptable profiles, however clean, is observation, not understanding — cap at 2.

> **Control strategy maturity.** Level 3 means the control strategy links critical process parameters to the quality attributes they drive (CPP→CQA), by design. Catching failures at final release testing, however reliable, is level 1–2 — testing quality in is not the same as designing it in. No documented parameter-to-attribute logic caps the score at 2.

> **Knowledge institutionalization.** If the honest answer depends on one or two named experts being reachable, the knowledge is tacit. It is a level 1 or 2 no matter how good those experts are — the instrument measures the system, not the individuals.

> **Transfer & network robustness.** If the process has never run at a second site or CDMO, you cannot claim it is transfer-proven. Cap at 2 until distributed execution has actually been demonstrated.

> **Regulatory commitment.** **(Exposure axis — magnitude, not evidence.)** Score by how hard regulatory positions have set: pre-filing and flexible is low; specs, methods, and comparability positions committed across filings is high. A high score is not a problem to fix — it is the recognition that a process change has become a regulatory event, not an engineering one.

05 — THE TRAPS, AND HOW TO BEAT THEM

Every self-assessment bends toward flattery in predictable ways.

Name the pattern in the room and the counter-move becomes obvious.

THE TRAP	HOW IT SOUNDS IN THE ROOM	COUNTER-MOVE
Success as readiness	*“We've never had a failed batch.”*	Ask what has actually been varied. Consistency at fixed conditions measures the width of your window, not the robustness of your process.
Confidence inflation	*“We're basically there.”*	Demand the artifact. With no data on the table, the score cannot exceed 2.
Seniority anchoring	*The most senior voice scores first and the room quietly follows.*	Score independently and silently before any discussion. Senior voices go last.
Optimism by omission	*Scoring the plan instead of the present.*	Score only what has been done. Planned characterization counts when it is complete — not when it is scheduled.
Middle-drift	*Everything lands at 2 or 3.*	Require at least one low and one high score to be defended with evidence. A uniform middle is usually avoidance, not calibration.
Gaming the zone	*“What do we need so we're not At the Wall?”*	The zone is an output, never a target. Score the dimensions honestly and let the zone fall where it falls.

06 — CROSS-FUNCTIONAL CALIBRATION

Run the calibration so the spread survives.

The calibration overlay captures the prism — the way one operational reality refracts into different interpretations across functions. It only works if functions score independently and privately first. The moment people discuss before scoring, they converge, and you destroy the very signal you are trying to read.

1. Each function places its read on overall readiness alone, without conferring.
2. Reveal all the scores at once.
3. Resist the urge to immediately resolve the spread. A wide spread does not mean someone is wrong; it means the organization has not yet aligned on reality. Ask each function why it sees what it sees. The competing explanations are the finding, not a problem to be smoothed over before the next agenda item.

07 — READING THE RESULT

Read the gap without flinching.

A high gap is information, not indictment. Read it plainly.

Understanding-Led / Intervention Window
The instrument is working exactly as intended. You have found a real gap while it is still cheap to close. Invest now.

Active Divergence
The window is closing. Treat this as a prompt to move characterization onto the critical path before it is forced there.

At the Wall / Beyond the Wall
Do not explain it away. The honest response is to name the specific exposures, assign owners, and shift from prevention to stabilization. A flattering re-score here is the most expensive mistake the tool can produce.

> **The failure is never a high score honestly reached. The failure is a low score that hides a real gap until the Wall makes it undeniable.**

08 — CADENCE

Score the trajectory, not the snapshot.

One score is a photograph; the framework's deeper claim is about rates. Re-score at every major stage gate and before each financing event, tech transfer, or manufacturing expansion — the same moments the white paper's executive questions are meant to interrupt. Keep the prior scores. A gap of 15 that was 5 two quarters ago is a program accelerating toward the Wall; a stable 15 is a managed position. Only the trajectory tells you which one you are.

FACILITATOR'S CHECKLIST

BEFORE

Sponsor frames the session as visibility, not judgment · PD/MSAT confirmed in the room · a neutral facilitator with no stake in the score is appointed.

DURING

Every dimension scored independently before discussion · the evidence test applied to every 3 or 4 (“what artifact defends this?”) · senior voices score last · at least one high and one low score defended · calibration scored privately, then revealed, then explored rather than resolved.

AFTER

The artifact behind each score recorded · the gap compared against last time to read trajectory · the top two or three exposures converted into owned actions before the next milestone.

On every score: “What would I show to defend this?”