ADVISORY REFERENCE · HAND IN GLOVE

Using Candid Cycle and Calibrate Capability together

How to get the most from the pairing for continuous improvement, with caveats and shorthand warning missives.

Candid Cycle tells you what people experienced.

Calibrate Capability tells you whether the organisation can deliver.

Each answers the question the other leaves open.

John Salter Consulting Services
disasterresilience.consulting · November 2026

PURPOSE

About this reference

This reference brings together a series of reflections on how Candid Cycle and Calibrate Capability can support continuous improvement together. It covers when to use them, how to sequence them, where the combination adds most value, and what limits it. It is written for sponsors, facilitators and assessors who want the pairing to change practice, not just produce reports.

The central proposition: Candid Cycle tells you what people experienced; Calibrate Capability tells you whether the organisation can actually deliver. Each answers the question the other leaves open.[1]

The central caution: the pairing structures improvement; it does not deliver it. Leadership, ownership and resources still do that.

At a glance

Section	What it covers
1 What each tool contributes	The voice, the evidence, and the weakness each has on its own
2 The joint loop	Four stages, two handoffs, and why the word “act” matters most
3 Where the synergy is strongest	Six ways the tools strengthen each other
4 When to use which	Matching the starting point to the situation
5 Practice principles	Ten principles for effective use
6 Caveats	What limits the pairing’s effectiveness
7 Shorthand warning missives	Fifteen quick reminders for the field
8 Signs the pairing is working	Four tests across successive cycles

SECTION 1

What each tool contributes

	Candid Cycle (the voice)	Calibrate Capability (the evidence)
Core question	What happened, what worked, and what did we learn?[1]	Is this capability established, applied in practice and embedded, and how far is it from where it needs to be?[1]
Input	Anonymous, structured reflections from contributors, who need no account or app[1]	Performance evidence across the seven Universal Framework domains[2,4]
Output	Recurring themes and follow-up actions[1]	Evidence strength (strong, mixed or weak), current and target maturity on the N/P/L/F scale, and improvement priorities[1,4]
Weakness on its own	Perception without proof	Proof without context

Calibrate tests each capability at three levels[2]:

Exists
The capability is documented in a policy, plan or procedure.

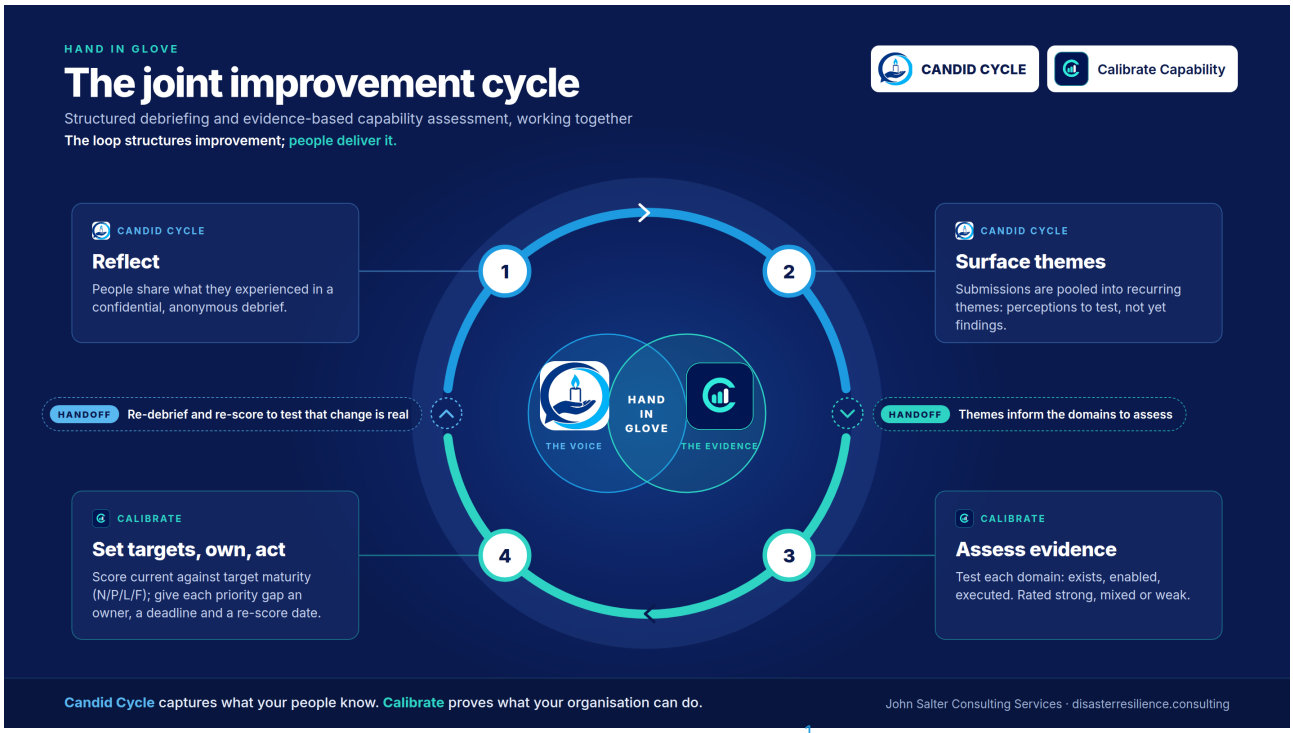
Enabled
The people, resources and authority are in place to make it work.

Executed
It is actually being done, and there is proof.

SECTION 2

The joint loop

The published model is a four-stage cycle, with Candid Cycle running the first two stages and Calibrate the last two.[1]



The joint improvement cycle, adapted from Hand in Glove with revised wording.[1]

1	**Reflect** Candid Cycle	The team gives its views in an anonymous debrief.
2	**Surface themes** Candid Cycle	Recurring themes emerge from those submissions: perceptions to test, not yet findings. **Handoff:** the themes inform which capability domains Calibrate assesses.
3	**Assess evidence** Calibrate	Each domain is tested for whether the capability exists, is enabled and is executed, and its evidence is rated strong, mixed or weak.
4	**Set targets, own, act** Calibrate	Current maturity is compared with target maturity, and each priority gap is given an owner, a deadline and a re-score date. **Handoff:** after action, a fresh debrief and a re-score test whether the change was real.

> **The word that matters most is "act".** Every other stage produces insight. Only this one produces change, and it is the stage the software can support but cannot carry out.

SECTION 3

Where the synergy is strongest

01 Themes set the scope
The assessment starts from what people raised, rather than from leadership's assumptions.[1]

02 Perception gets tested
A theme such as "decisions aren't communicated" is checked against documents, practice and outcomes, so a genuine capability gap can be told apart from a one-off frustration.[1]

03 Evidence gets context
A weak rating shows where a gap is; debrief reflections help explain why it exists and how it is felt on the ground.[1]

04 Change gets verified twice
Reassessment shows whether capability has improved; a later debrief shows whether people experience the difference. (Advisory interpretation of the published loop.)

05 Reviews get follow-through
After a culture, post-incident or post-election review, the pairing helps sustain change instead of leaving the report on a shelf.[1]

06 Governance is consistent
Both tools use AI-drafted reports that the account holder reviews before release.[1]

The distinctive value is that the pairing joins **organisational learning** ("What did we learn?") to **capability assurance** ("What can we demonstrate?") and connects both over time.

SECTION 4

When to use which

Situation	Suggested approach
After an event, project or incident, with no clear sense of the issues	Start with Candid Cycle; bring in Calibrate if themes point to systemic gaps
Board or regulator needs evidence of maturity	Start with Calibrate; use Candid Cycle to test whether ratings match lived experience
Following an independent review	Use both: debrief to capture response, assess to baseline, repeat to show progress
A single, contained concern	A personal Candid Cycle review or a Calibrate rapid review of one domain[1]
Organisation-wide improvement programme	Multiple team cycles plus a full seven-domain assessment, repeated against the same criteria[1,2]

The practice's own answer to "which one?" is "yes, no, or maybe, depending on your needs and purpose, and sometimes both".[5]

SECTION 5

Practice principles for effective use

1. **Candid first, Calibrate second, as a default rather than a rule.** Let themes guide scope, but always include governance-critical domains even if nobody raised them.
2. **Seek independent evidence.** Where possible, the evidence Calibrate assesses should not come solely from the people whose accounts it is testing.
3. **Rate evidence honestly.** "Mixed" and "weak" are useful findings. Findings based on attestation are flagged as not independently verified[2]; keep that distinction visible to decision-makers.
4. **Interpret before prescribing.** Consider alternative explanations and contrary evidence before turning a theme into an intervention.
5. **Give every priority an owner, a deadline and a re-score date.** Without them, findings remain a report.[3]
6. **Keep one improvement backlog.** Debrief actions and assessment priorities should merge, not compete.
7. **Reassess against the same criteria.** Set a baseline, act, then measure change on a like-for-like basis.[2]
8. **Review the AI-drafted reports properly.** The account holder's review is a governance control, not a formality.
9. **Close the loop with contributors.** Tell people what changed because of what they said; it sustains future candour.
10. **Keep the effort proportionate.** Match the depth of the cycle to the significance of the issue.

SECTION 6

Caveats

Shared blind spots

If the same people frame the questions, provide the accounts, select the evidence and judge the results, the second tool may confirm the first rather than test it. The Exists / Enabled / Executed distinction is the main safeguard, but only if it is applied rigorously.

Candour is not guaranteed

Anonymity supports candour; it does not prove it. Check who was invited, who responded, who was missing (contractors, frontline staff, partners, affected clients) and whether people believe their input will make a difference.

Translation requires judgement

“Decisions took too long” might reflect unclear authority, missing information, competing priorities, inadequate skills, or appropriate caution. The pairing surfaces the issue; skilled interpretation diagnoses it.

Coverage can become too narrow

Debriefs look back at what people have experienced. Some critical capabilities have not been tested recently. The seven-domain framework is a structure for inquiry, not proof that every organisation-specific exposure has been examined.

Insight is not implementation capacity

The pairing does not supply budget, authority, time, skills or incentives. An organisation can run both tools competently and still make little progress.

Better scores are not necessarily better outcomes

Distinguish three tests:

Test	What it establishes
Action completed?	The agreed task was done
Capability improved?	A stronger way of working can be demonstrated
Outcomes improved?	The change produced a worthwhile practical result

Two processes can create friction

Re-entering data, parallel action lists or unclear handoffs can turn synergy into overhead. Confirm that a theme can be traced to a finding and then to an action.

Scope of explanation

Both tools assess organisational capability and experience. They do not explain external factors, such as why voters, customers or the market behaved as they did.[3]

Evidence of impact

The published materials explain the mechanism well. Comparative outcome evidence showing that the pairing reliably produces sustained improvement has not yet been established. Claims should be framed accordingly.

SECTION 7

Shorthand warning missives

Short reminders for facilitators, sponsors and assessors. Suitable for a briefing card, a slide footer or the top of a facilitation plan.

01 Not a silver bullet.
It structures improvement; people deliver it.

02 Heard is not proven.
A theme is a perception until evidence tests it.

03 Documented is not done.
Exists is not Executed.

04 Anonymous is not automatically candid.
Check who spoke and who did not.

05 Silence is not safety.
Unraised domains may still be critical.

06 Correlation is not cause.
Interpret before you prescribe.

07 No owner, no outcome.
Every priority needs a name and a date.

08 Done is not better.
Completed actions are not improved capability.

09 Better scores are not better results.
Check the outcome too.

10 Attestation is not verification.
Keep the flag visible.

11 Review the draft.
AI drafts; accountable people decide.

12 One list, not two.
Merge the backlogs.

13 Close the loop.
Tell contributors what changed.

14 Same ruler, every time.
Reassess against the same criteria.

> **15 Insight without action is just a better-written shelf report.**

SECTION 8

Signs the pairing is working

Across successive cycles, ask:

- Did the pairing reveal important gaps that either tool alone would have missed?
- Did findings lead to implemented changes, not just recommendations?
- Did reassessment show stronger capability, corroborated by later debriefs?
- Did relevant operational outcomes improve and stay improved?

Consistent “yes” answers are the evidence needed to move from a strong design argument to a demonstrated impact claim.

Positioning line

> **Candid Cycle and Calibrate Capability work hand in glove to turn experience into evidence-informed action, and to test whether that action makes a lasting difference.**

SOURCES

Sources

1. John Salter Consulting Services, Hand in Glove: Candid Cycle and Calibrate Capability (7 November 2026). https://disasterresilience.consulting/2026/11/07/hand-in-glove-candid-cycle-and-calibrate-capability/
2. John Salter Consulting Services, Blog. https://disasterresilience.consulting/blog-2/
3. John Salter Consulting Services, Separate what you controlled from what the environment gave or took away (6 November 2026). https://disasterresilience.consulting/2026/11/06/separate-what-you-controlled-from-what-the-environment-gave-or-took-away/
4. Calibrate Capability, Apple App Store. https://apps.apple.com/us/app/calibrate-capability/id6759160398
5. John Salter Consulting Services, home page. https://disasterresilience.consulting/

Points not marked with a source number are advisory interpretation by the author. This reference is general guidance and is not a substitute for professional judgement in a particular context.