New Zealand Journal of Health and Safety Practice
He rangi ta matawhaiti
He rangi ta matawhanui

A person with narrow vision has a restricted horizon; a person with wide vision has plentiful opportunities

Protecting the Protectors: Beyond Verification to Critical Control Reliability, Controls protect people. People protect controls

Vol 3; edition 4; November, 2026

Author: Deborah Cameron

Email: Deborah@redhawksafety.nz

DOI: https://doi.org/10.26686/nzjhsp.v3i4.11809

Abstract

Critical Control Management (CCM) has significantly improved organisations' ability to identify, verify and monitor controls that prevent fatal and serious incidents. Yet verification only confirms that a control was fit for purpose on the day it was checked; it says nothing about whether that control will still be relied upon and maintained by workers between formal checks.

This paper introduces Care-Based Critical Control Reliability (CBCCR), a construct that locates the reliability of a critical control in the organisational conditions that motivate and enable people to actively maintain, challenge and strengthen it in everyday practice. The accompanying Care Activation Model (CAM) sequences these conditions: Organisational Care, Connection, Trust and Psychological Safety, Collaboration and Shared Responsibility, and Intervention Behaviours, culminating in Critical Control Reliability. Connection is proposed as a precluding, previously unnamed, gateway condition without which care does not translate into trust or voice.

Two anonymous case studies illustrate how technically identical, fully verified controls diverged in practice depending on this relational foundation. CBCCR is offered as a bridging mechanism between the CCM literature and existing theories of trust, psychological safety and organisational drift.

Keywords
critical control management; organisational care; connection; psychological safety; intervention behaviours; critical control reliability

1.0 Introduction

Becoming a Certified Fellow made me look harder at something that had never quite sat right: when I sign off that a critical control is verified, what am I actually claiming? Only that it exists, and that someone checked it, on the day they checked it. I cannot claim it will still be doing its job tomorrow, under production pressure, on a different shift, when no safety professional is standing nearby.

Critical Control Management (CCM) (International Council on Mining and Metals [ICMM], 2026) has become very good at assuring the integrity of barriers that protect people, but has under-theorised the integrity of the humans who sustain those barriers. This raises a question rarely asked directly: what protects the controls themselves? A permit-to-work system, a machine guard, an isolation procedure - each is engineering or documentation, but each is also embedded in a living workplace, sustained or eroded by the people who use it every day.

This gap between verification and reliability is not new. CCM has genuinely improved organisations' ability to identify, verify and monitor controls that prevent fatal and serious incidents. Yet the pattern persists across industries: technically sound controls degrade, are worked around, or quietly stop being followed, despite passing every formal assurance check along the way. Traditional CCM asks whether a control exists and has been verified. This paper asks a different question: what organisational conditions are required for a control to remain reliable between verifications, when no one is watching?

Why this matters now. New Zealand's Health and Safety at Work Act (HSWA) has been amended, with changes taking effect from 1 April 2027, introducing HSWA's first statutory definition of “critical risk” and requiring officers to verify that resources and processes for managing it are being used

(Health and Safety at Work Amendment Act 2026; WorkSafe New Zealand, 2026). That word, verifying, is where this paper begins and departs. The law will sharpen assurance that critical risk controls exist. It cannot tell an officer whether a control will still be holding on the night shift, six months after the last audit, when it is quietly easier not to bother. Legislating a focus on critical risk raises the stakes of verification; it does not resolve what keeps a control reliable in between.

This paper is written from, and for, a New Zealand context, and acknowledges Mātauranga Māori and the contribution of Te Ao Māori concepts to workplace health and safety practice. In particular, manaakitanga (caring for and supporting others) and whanaungatanga (relationship, connection and a sense of belonging) resonate closely with the Organisational Care and Connection constructs developed in this paper, and are acknowledged as part of the wider cultural foundation from which they draw meaning in a New Zealand workplace.

This paper argues that critical controls do not fail in isolation - they fail in context. The research question: what organisational conditions enable frontline workers to actively maintain, challenge and strengthen critical controls in day-to-day practice, rather than relying on point-in-time verification alone? The next section sets out why controls drift and fail even when technically sound, before introducing the Care Activation Model - a framework for what sustains critical controls, and, in turn, the people they protect.

2.0 Introducing Care-Based Critical Control Reliability

I want to give this gap a name, because until a problem has a name it is hard to build a discipline around it. I call it Care-Based Critical Control Reliability (CBCCR): the reliability of a critical control depends not only on its technical design and verification history, but on the organisational conditions that motivate and enable people to actively maintain, challenge and strengthen that control in day-to-day practice. Verification tells you a control was fit for purpose on the day it was checked. CBCCR is concerned with what happens in all the days in between.

A critical control is usually described and assured in technical language - specification met, inspection passed, certificate current. But no control operates on its own; it is picked up, interpreted, and sometimes defended by people working under real constraints: fatigue, time pressure, competing priorities. A control is better understood as a socio-technical system, sustained by human behaviour, not despite it. The engineered control and the procedure don't drift on their own; the behaviour around them does.

Controls that are technically sound and formally verified still fail, in a small number of well-documented ways:

- Drift into failure - small, individually reasonable deviations accumulate unnoticed (Dekker, 2014).
- Normalisation of deviance - a harmless deviation becomes the new “normal,” luck mistaken for evidence (Vaughan, 1996).
- Workarounds under production pressure - when the control is slower than the job demands, people solve the conflict themselves.
- Failure to challenge emerging weaknesses - someone notices, but doesn't say anything. The gap isn't awareness, it's voice.

None of these require a technically deficient control, and none would necessarily be caught by a scheduled verification. They are behavioural, cultural, and entirely predictable - which is exactly why they deserve a framework, not a hope.

People are also the primary mechanism that can stop that degradation - but whether someone intervenes depends less on training than on the organisational environment they're standing in at the moment it matters: whether raising a concern is treated as diligence or friction, whether a missed step is met with curiosity or blame. A frontline worker deciding whether to intervene is, in effect, reading the organisation's leadership, not the procedure alone. This is the real question CBCCR is trying to answer: not whether a control exists, but whether the organisation around it is one where people are motivated and enabled to keep it alive.

3.0 How Do We Achieve This? The Care Activation Model

If people are the stabilising mechanism against drift, normalisation, workarounds and silence, the practical question becomes: what has to be true of an organisation before that stabilising behaviour

actually shows up? The Care Activation Model (CAM) answers this as a sequence, not a checklist: Organisational Care, Connection, Trust and Psychological Safety, Collaboration and Shared Responsibility, Intervention Behaviours, Critical Control Reliability. Each stage depends on the one before it holding. The model's central claim, and its point of departure from existing literature, is that this sequence has a gate in it: Connection. Skip it, and everything downstream - trust, voice, intervention - is like wiring a system without closing the circuit (see Figure 1).

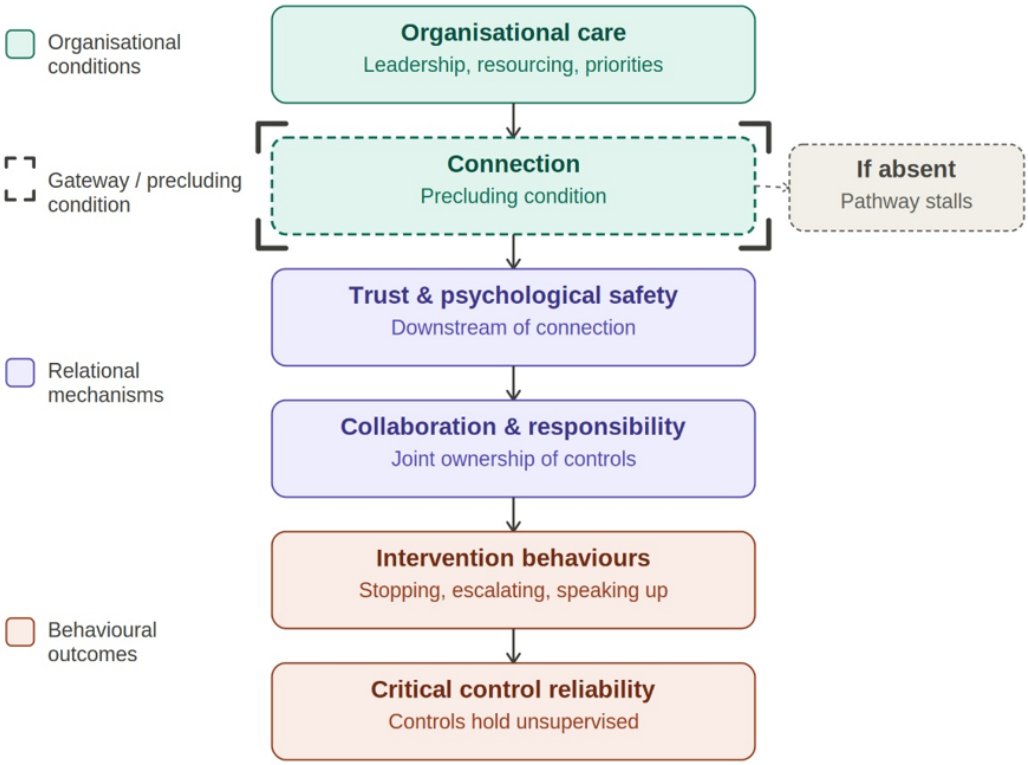


Figure 1. The Care Activation Model (CAM).
Organisational conditions (teal) enable connection, which is shown with a dashed border to indicate it is a precluding condition: without it, the pathway to trust and psychological safety stalls. Connection enables the relational mechanisms (purple), which in turn enable the behavioural outcomes (coral) that sustain critical control reliability.

3.1 Organisational Care

I use “care” in a deliberately unsentimental sense. Organisational Care is not warmth, and not an HR value on a poster. It is a systemic capability, demonstrated through leadership decisions, resourcing choices and operational priorities that signal genuine commitment to people over short-term convenience - the roster that isn't quietly understaffed, the leader who changes a production schedule rather than let a control lapse under time pressure. Care is something an organisation does, visibly and repeatedly, not something it feels.

3.2 Connection — the precluding step

This is the model's novel move: Organisational Care does not translate into Trust directly. It must pass through Connection first. Connection is the felt, relational experience of being seen, known and valued by leaders and peers - a human-to-human bond, not a policy. Without it, even genuine care initiatives read as transactional or performative, and trust does not take root, no matter how sincere the intent.

Connection functions as a precluding condition rather than a contributing one. If it's absent, the rest of the pathway stalls entirely. Care can't be mandated, but Connection has practical, structural antecedents an organisation can influence: team stability, span of control, rostering continuity, and leader floor-time. These are levers, not guarantees; operationalising how they interact is flagged as a direction for further research.

3.3 Trust, Psychological Safety, and Why Frontline Workers Speak Up

Edmondson's (1999) work explains why people stay silent about risk even when they've noticed it. Where CAM departs is in sequencing: Edmondson treats psychological safety largely as an endpoint leaders build directly. CAM treats it as contingent on something prior - it cannot be built onto a workforce that hasn't first experienced Connection. A leader can practise every psychological-safety behaviour in the literature and still not get voice from a team that doesn't feel known by them.

This matters for critical controls because so many of the moments that count are small and ambiguous - a permit that looks slightly wrong, a guard propped for “just this job.” Stopping unsafe work or challenging an inadequate verification requires a worker to believe, in that moment, that speaking up will be received as diligence rather than disruption. Sincerity and training alone don't produce that belief; both sit on a relational foundation that either exists or doesn't.

3.4 Intervention Behaviours

This is where the model cashes out into observable behaviour. Where Care, Connection, Trust and Psychological Safety are present, workers will:

- Stop unsafe work, even when it costs time or invites friction
- Notice and challenge an inadequate verification or absent control, rather than assume someone else has checked
- Escalate emerging control weaknesses before they become failures
- Speak up when risk is increasing, even when the signal is ambiguous

These behaviours are how Critical Control Reliability actually gets produced, day to day. Critical Control Reliability doesn't come from workers being told to do these things - it comes from workers who feel and know they matter enough for it to be worth doing. Training gives workers the skill to intervene; Care, Connection and Trust give them a reason to use it.

4.0 A Contemporary Critical Control: Permit to Work Systems

It's worth grounding these principles in a control every safety professional will recognise: the Permit to Work (PTW) system — often treated as the archetypal “hard” control, but unavoidably a socio-technical one too.

Caponecchia and Wyatt (2021) and Braun et al. (2021) argue that a genuinely safe system of work is broader than any single task procedure - it depends on a set of conditions holding together. Applied to a PTW, the necessary technical conditions are competent workers, suitable plant and environment, and ongoing monitoring - the layer CCM verification is good at checking. A Manager Assurance Check can confirm each condition was true at a point in time, every box honestly ticked. What none of those conditions tell you is what happens the moment conditions change after sign-off - when almost every PTW-related incident actually occurs. CBCCR adds a relational necessary condition alongside the technical ones: whether the permit-holder believes stopping work is something they can do, without cost. It isn't on any audit checklist because it isn't a property of the permit - it's a property of the relationship between worker and organisation, built, or not, through the Care to Connection to Trust pathway.

Consider a permit issued correctly - isolation confirmed, competencies checked, every signature in place. Partway through, conditions shift: a nearby process starts unexpectedly, or the isolation window is closing. Nothing on the permit itself is wrong. The worker now faces a decision no document can make for them: stop and accept the friction of reopening a signed-off permit, or continue on the reasonable-sounding basis that the change probably doesn't matter. An audit the following week will find a correctly signed permit. It will not find, and cannot find, whether the worker who held it felt able to stop - which is the gap between assurance and reliability this paper is built around.

5.0 Bringing in Experience: Small-Scale Case Studies

Frameworks are only as useful as the practice they explain. The two cases below are de-identified composites from my own experience: one where a control quietly failed despite being verified, one where an intervention actually happened.

Table 1. Two cases

Case	Brief background	CBCCR / CAM lens: which link broke down?
A	A guard-interlocked machine stopped the line for manual ejection of defective product. Under sustained production pressure and an unchanged throughput target, the interlock was routinely bridged.	The interlock had passed verification repeatedly. No one had reconciled the production target with the control, and the team had no trusted channel to raise the mismatch without it reading as a complaint — the break was at Organisational Care/Connection.
B	A working-at-height task required a High Risk Work Permit specifying a Safety Observer. The observer was pulled onto another task after start-up, and work continued without one - until a contractor noticed, stopped the job, and refused to continue until someone was back in position.	The permit was correctly issued. The gap between paper and practice was closed by Intervention Behaviour, built on a prior history of leaders responding to concerns with curiosity rather than irritation - Connection and Trust, established well before this job.

Case A: a scheduled CCM verification would likely have recorded “interlock in place, tested, functioning,” but would not have captured the control being bridged during ordinary production. CBCCR reveals the real mechanism: a target and a control never reconciled, in an environment where raising the mismatch didn't feel safe.

Case B: a permit audit would also have shown a compliant picture. It would not have shown that the named observer was absent at the point it mattered. What closed that gap was one worker's willingness to stop a job that wasn't theirs to authorise, on relational confidence that doing so wouldn't cost them.

Both controls were, on paper, in exactly the same state — verified, current, documented. Whether either one actually held came down to something verification doesn't measure: a relational basis to raise a mismatch, absent in Case A, or to intervene in the moment, present in Case B.

6.0 Discussion: How Do We Gain Assurance Controls Will Actually Work?

Every Senior Leadership Team eventually asks some version of the same question: have critical controls been verified? It's the right question for CCM's original purpose, and the wrong one for this paper's purpose. Alongside it belongs another: are people actively protecting critical controls - and do they feel enough connection to do so? The first can be answered by an audit. The second cannot, which is precisely why it has been missing from most assurance frameworks.

“Connection” and “care” can sound unmeasurable next to a verification checklist, but the indicators simply look different, and lead rather than lag: control-related intervention rates, escalation and challenge frequency, psychological safety pulse measures, and relational indicators such as leader–worker contact around control performance. None of this requires new technology; most organisations already collect adjacent data. The shift CBCCR asks for is treating it as a leading indicator of control reliability, reported alongside verification data rather than buried in a separate wellbeing report.

6.1 Positioning Against Existing Theory

This paper does not claim to have discovered trust, psychological safety, or speaking-up behaviour - those are well established. What CBCCR offers is narrower: a bridging mechanism between two literatures that currently run in parallel.

- Reason's (1997) Swiss Cheese Model explains how accidents happen once defences fail, not why people do or don't intervene beforehand.
- Dekker (2014) and Vaughan (1996) explain drift into failure and normalisation of deviance with precision, but not what sustains resistance to it.
- Edmondson (1999) explains speaking-up behaviour, but treats psychological safety as an endpoint leaders build directly, not as contingent on Connection (see 3.3).
- Caponecchia and Wyatt (2021) and Braun et al. (2021) broaden “safe system of work” beyond a single task procedure - useful company for the claim that controls are socio-technical systems.
- Hummerdal (2026) argues that organisations that cannot hear themselves cannot improve themselves, positioning conversation as the mechanism through which weak signals travel and

frontline intelligence reaches leadership - closely aligning with CBCCR's proposition that Connection precedes trust, voice and intervention behaviours.

CCM treats culture and human behaviour as a parallel track alongside technical verification, with no theorised mechanism connecting the two. CBCCR is offered as that mechanism, with Connection as its precluding, previously unnamed, first step - a hypothesised antecedent variable requiring empirical validation, not an established causal relationship.

7.0 Conclusions

This paper began with a question I couldn't ignore once I looked at it properly: when I sign off that a critical control is verified, what am I actually claiming? Only that it worked on the day someone checked it - not that it will still be holding tomorrow, on the night shift, when no one is watching.

This paper makes two contributions. First, conceptually: naming Care-Based Critical Control Reliability (CBCCR) as a construct distinct from CCM, locating reliability in the organisational conditions that determine whether people sustain a control between verifications. Second, practically: the Care Activation Model, sequencing Organisational Care, Connection, Trust and Psychological Safety, Collaboration, and Intervention Behaviours toward Critical Control Reliability — with Connection positioned as a precluding, foundational step.

This matters for practitioners because the law is about to get sharper on half the problem. From 1 April 2027, HSWA will require officers to verify that resources and processes for managing critical risk are being used (Health and Safety at Work Amendment Act 2026). That improves assurance that controls exist. It does nothing to ensure people feel connected enough to protect them when it counts - which remains a leadership responsibility, not a legislative one. It matters for scholarship because it bridges two literatures that have run in parallel until now. The legislation strengthens verification. CBCCR seeks to strengthen reliability.

This is a starting framework, not a finished one. The clear next step is operationalising Connection itself, and testing in a follow-up paper whether stronger Connection lead indicators predict higher intervention rates and fewer quiet control failures.

Controls protect people. This paper's claim is that people protect controls - and whether they do depends on whether they feel connected enough to bother. Verification tells you a control was fine on the day you looked. Connection tells you whether it will still be fine tomorrow, when you don't. Controls protect people. People protect controls.

Declarations

Funding sources: Self-funded by the author.

Conflicts of interest: Declarations of interest: none.

Disclaimer: The views expressed in this paper are those of the author and do not represent the views or positions of any current or former employer.

Use of Artificial Intelligence: Generative AI (Claude, Anthropic) was used during manuscript preparation to assist with language refinement and copyediting - identifying repetitive phrasing and stylistic patterns for revision. The conceptual framework (CBCCR, the Care Activation Model), case material, argument, and analysis are the author's own original work. AI was not used to generate ideas, content, data, or citations. The author reviewed, edited, and takes full responsibility for the final manuscript.

8.0 References

Braun, C., Brophy, S., Jassim, M., Mills, E., Dagar, P., Peace, C., Rose, A., & Wichman, M. (2021). A structured approach to the concept of a (safe) system of work. New Zealand Journal of Employment Relations, 46(2), 4–19.

Caponecchia, C., & Wyatt, A. (2021). Defining a safe system of work. Safety and Health at Work, 12(4), 421–423. https://doi.org/10.1016/j.shaw.2021.07.001

Dekker, S. (2014). The field guide to understanding human error (3rd ed.). CRC Press.

Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. Administrative Science Quarterly, 44(2), 350–383. https://doi.org/10.2307/2666999

Health and Safety at Work Amendment Act 2026 (New Zealand).

Hummerdal, D. (2026). An invitation to safety conversations. Open Change Press.

International Council on Mining and Metals. (2026). Critical control management: Good practice guide. ICMM. https://www.icmm.com/en-gb/guidance/health-safety/2026/ccm-good-practice-guide

Reason, J. (1997). Managing the risks of organizational accidents. Ashgate Publishing.

Vaughan, D. (1996). The Challenger launch decision: Risky technology, culture, and deviance at NASA. University of Chicago Press.

WorkSafe New Zealand. (2026). Understanding the changes to the health and safety law. https://www.worksafe.govt.nz/laws-and-regulations/acts/changes/understanding-the-changes-to-health-and-safety-law/