ARM POLICY LAB | CONTROLLED PROTOCOL

VERSION 3.0



GOVERNANCE RESILIENCE FRAMEWORK GRF v3.0

Evidence, corrective actions, institutional learning and economic resilience

CONSOLIDATED RESEARCH PROTOCOL

AUTHOR

Armin Mešić, CPA

Continuing GRF v2.2 | 12 October 2026

© 2026 ARMIN MEŠIĆ, CPA. ALL RIGHTS RESERVED.

Research framework without confirmed empirical validation.

Purpose and publication status

GRF v3.0 sets out how organisational governance capability is assessed through verifiable evidence of signal preservation, reasoned decisions, corrective actions and learning. It continues GRF v2.2 of August 2026 and consolidates later methodological changes into one operational text. Its primary purpose is organisational diagnosis and research; empirical validation has not been confirmed.

This is an editorially finalised publication candidate in English, translated from the consolidated Bosnian edition. Editorial completion does not mean that independent methodological review G5 or the record of approval and rights G6 has been completed. In the available source package, these controls remain undocumented. The edition therefore does not carry unconditional publication approval.

Field	Control value
Author	Armin Mešić, CPA | ARM Policy Lab
Version and date	3.0 | 12 October 2026
Predecessor	GRF v2.2 | August 2026
Development inputs	Materials v2.2.2 and v2.3 and protocol v3.0-RC1; labels do not establish separate public release.
Validation status	Proposed research and diagnostic framework; no confirmed empirical validation.
Release status	Conditional: G5 and G6 remain open according to the available record.
Case supplement	Questions to ArcelorMittal Group are separate from this protocol.

Copyright and limits of use

© 2026 Armin Mešić, CPA. All rights reserved in the original authorial expression and GRF materials to the extent provided by applicable law. Rights in third-party sources remain with their holders. Lawful quotation and other exceptions are unaffected. Recommended citation: Mešić, Armin, CPA. Governance Resilience Framework GRF v3.0 — Evidence, corrective actions, institutional learning and economic resilience. ARM Policy Lab, 2026. English edition.

GRF is not a certification standard, audit opinion, legal determination, credit or investment rating, or bankruptcy prediction model. It does not establish intent, liability or causation. It must not be presented as a product, validation or endorsement of ISO, OECD or another institution. Third-party MEAID and EDRM instruments are not included as modules with resolved rights.

Contents

1 Conceptual architecture and transition tests
2 Evidence discipline and indicator assessment
3 Measurement model and coverage rules
4 Manual of 16 capability indicators
5 Exposure, controls and independent outcomes
6 Empirical research and reliability
7 Registers, evidence integration and implementation
8 Release controls and sources

1 Conceptual architecture and transition tests

Formal policies constitute evidence of design. Effectiveness requires operational implementation, a record of evidence consideration, an accountable decision, action verification and preserved learning. An adverse economic outcome may also arise under strong governance when dependencies or external shocks are high. Capability, exposure, controls, outcomes and context are therefore measured separately.

Layer	Construct	Question and boundary
A	GRCS	Can the organisation detect, preserve, reason, correct and learn? An outcome does not replace capability evidence.
B1	ITE	How high is inherent dependency before mitigation? High exposure is not misconduct.
B2	ECE	How effective are tested continuity controls? Policy is not evidence of effect.
B3	RTE	Experimental index of remaining exposure. It is not interpreted as probability or expected loss.
C	Outcomes	What occurred after the baseline cutoff? This does not change the earlier rating.
D	Context	Sector, size, ownership, regulation, geographic and macroeconomic conditions. This is not a rating.

Five principles

Traceability and institutional memory: material signals, decisions and reasons must remain reconstructable by authorised users.

Protected reporting and signal preservation: reporting must be safe, evidence preserved and handling impartial.

Learning and adaptation: a finding acquires organisational value through accountable change and verification of effect.

Governance adaptability and trust: trust requires consistent consideration of evidence and proportionate response.

Economic continuity: employee, supplier, infrastructure and community dependencies must be visible in material decisions.

The principles are cumulative. Transparency without protection can expose the reporting person; protection without action can result in procedural legitimacy alone; learning without a responsible owner can remain local and disappear during change.

GRTM transmission model

Stage	Proposed mechanism
1 Governance weakness	Weak oversight, procedural ambiguity or an unresolved control deficiency.
2 Loss of learning	Warnings, findings and incidents do not become available institutional knowledge.
3 Reduced adaptability	The response remains reactive, temporary or too narrow.
4 Declining trust	Stakeholders respond to uncertainty by delay, protective action or exit.
5 Economic transmission	Changes in continuity, payments, employment, supplier relationships or investment.
6 Local vulnerability	Where dependency is material, effects may weaken local diversity and resilience.

Every transition is a hypothesis to be tested. The assessor records sequence, observable mechanism, alternative explanations, break conditions and evidence quality. The model allows feedback: liquidity pressure, staff loss or supplier disruption can affect governance.

Operational transition tests

Transition	What is checked	Break condition
Signal → preservation	Acknowledgement, date, protection and evidence provenance.	Loss, fragmentation or unavailable provenance.
Preservation → classification	Separate facts, allegations, risks and relevant standards.	Premature narrowing or concealed ambiguity.
Classification → escalation	Materiality, authority, competence and independence.	Conflict of interest or unexplained non-escalation.
Escalation → investigation	Mandate, scope, access and participation.	Selective evidence or a non-reconstructable trail.
Investigation → decision	Material evidence and competing explanations.	A conclusion without links to evidence and reasons.
Decision → action	Individual and systemic effects, responsibility and deadline.	Closure without reasoned assessment of action.
Action → effect	Retesting and recurrence monitoring.	Activity completion without evidence of effect.
Effect → learning	Preserved lessons, change ownership and adaptation.	Isolation or deletion of the earlier signal.

Three separate forms of case closure

Closure	Minimum evidence	What it does not establish
Procedural	The process is formally completed and its status recorded.	That harm has been remedied or the decision is evidentially complete.
Corrective	Individual and systemic effects considered; action implemented or declined with reasons.	That the action is sustainably effective or learning has occurred.
Learning closure	Cause and recurrence translated into accountable changes, retesting and preserved lessons.	That every request has been accepted or every disputed fact resolved.

During ownership change, open cases, preserved and transferred records, legal and operational holders of obligations, deadlines and retesting are recorded separately. Transaction motive, assumption of liability or legal succession must not be inferred from event sequence.

2 Evidence discipline and indicator assessment

Before scoring, freeze the question, organisational level, period, information cutoff, applicable indicators and hypotheses. The group, parent company, local legal entity and individual transaction have separate evidential boundaries.

Chain of epistemic traceability

Procedural event → evidence → source classification → factual claim → relevant standard → reasoned conclusion → decision → action or reasoned non-action → monitoring → learning record. A gap in the chain remains an evidence gap; it does not itself establish misconduct, intent or non-functionality.

Status	Meaning	Boundary
CONFIRMED	A bounded claim directly established by an adequate primary or official record.	Does not automatically establish liability or causation.
CORROBORATED	A materially separate information pathway supports the same claim.	Repetition of the same original allegation is not corroboration.
NOT ESTABLISHED	The sufficiency threshold has not been met.	Does not establish that the opposite claim is true.
CONTRADICTED	Reliable evidence affirmatively challenges the claim.	Does not automatically establish intent.
MATERIALIZED RISK	A later outcome corresponds to a previously recorded risk.	Does not feed the outcome into the baseline rating or establish cause.

Source classes are primary event evidence, official procedural record, organisational operational record, party submission, independent external verification, analytical synthesis and public context. Sending a letter establishes dispatch and the content of its allegations. Receipt, acceptance, action and institutional endorsement require separate records. Author-produced matrices are not independent corroboration of the author’s own claims.

Separate assessment axes

Axis	Codes and rule
Performance	P0–P4: ordinal functioning anchors; do not treat them as proven equal intervals.
Evidence quality	E0 assertion; E1 design/policy; E2 operational application; E3 independent review or retest; E4 longitudinal learning/adaptation.
Assessment confidence	Low, medium or high; explain provenance, independence, completeness, temporal alignment and reproducibility separately.
Applicability and missingness	NA is non-applicability justified in advance. NE-ACCESS indicates unavailable evidence; NE-QUALITY indicates insufficient quality.
Claim status	The evidential status of a claim is separate from P, E and confidence.
Red flags	RF0–RF3 are recorded separately and must not be concealed by an average.

The P0 and NE rule

P0 requires affirmative evidence that a practice is absent or materially non-functional. Absence of an available file, policy or measurement is insufficient. If the evidence does not meet the threshold, the rating is NE. This rule replaces earlier anchors that equated missing evidence or lack of measurement with P0, particularly for SPK-03.

Rating	General anchor
P0	Demonstrably absent or materially non-functional practice.
P1	Reactive, inconsistent application, predominantly after failure.
P2	Defined responsibilities and processes supported by operational evidence.
P3	Comparison over time, accountable changes and verification of effect.
P4	Adaptive governance, appropriate independent verification and sustained effect.
NE	Unavailable or insufficient evidence; no P rating is assigned.

Red flags

Code	Treatment
RF0	No qualifying flag identified; this does not mean that risk is absent.
RF1	A credible signal requires targeted review.
RF2	Material integrity, protection, action or continuity concern; disclose alongside results.
RF3	Critical threat to evidence, independence, protection or basic continuity; the composite must not be presented alone.

Typical signals include altered evidence, unclear action ownership, unexplained exclusion of contrary evidence, adverse treatment after reporting without review, recurrence after closure

and loss of records during a transaction. RF must not be assigned on the basis of reputation, company size, outcome or unavailability without further justification.

3 Measurement model and coverage rules

The four capability domains are GOV, CMP, SPK and LRN, each containing four indicators. Only indicators with E2–E4 evidence supporting the assigned anchor enter the composite. E1 alone does not justify an operational rating. Maturity is not an automatic function of an average.

For domain d: Dd = mean(Pi / 4) across eligible indicators in that domain. GRCS = 100 × (0.25 × DGOV + 0.25 × DCMP + 0.25 × DSPK + 0.25 × DLRN). Equal weights are provisional research assumptions, not empirically established contributions.

ITE = 100 × mean(ITEj / 4). ECE = 100 × mean(ECEj / 4). Experimental RTE = ITE × (1 − ECE / 100). ITE increases with inherent dependency, whereas ECE increases with control effectiveness. P0 for ITE denotes low exposure; P0 for ECE requires demonstrated control non-functionality. The directions must not be reversed.

ECE is not a validated percentage of losses prevented. RTE is not a probability of harm, an actuarial expected loss or a proven proportion of risk reduction. Separate ITE and ECE results are the primary presentation. Research compares aggregate and paired models, weights, missingness and sector sensitivity.

Result	Condition for presentation
GRCS	Overall eligible coverage ≥75%, and E2+ evidence for ≥50% of applicable indicators in each domain.
ITE and ECE	Present separately with counts of applicable and scoreable records, NE and limitations.
RTE	Both constructs have coverage ≥75%; at least half the records in each paired category must be scoreable.
Threshold not met	Withhold the composite; present individual evidence, coverage and the reason.
NA and NE	Exclude NA from the applicable denominator; retain NE in the denominator without converting it to zero.

Coverage = number of indicators with an eligible rating / number of indicators deemed applicable in advance. If all 16 indicators apply, at least 12 must be scoreable and at least two in each domain. If four ITE and four ECE indicators apply, at least three are required in each construct. For RTE, also check every pair; a wholly unavailable member of a pair fails the category condition.

Formula illustration: if the four domain means of P ratings are 3, 2, 3 and 2, the normalised domains are 0.75, 0.50, 0.75 and 0.50, giving GRCS = 62.5. If ITE = 75 and ECE = 50, RTE = 37.5. These values are a synthetic calculation example without organisational or empirical interpretation: SYNTHETIC_TEST_FIXTURE=YES.

GRMM maturity

Level	Evidential basis
1 Fragmented governance	Reactive and fragmented roles, trails and follow-up.
2 Structured governance	Operationally defined responsibilities, procedures and controls.
3 Learning governance	Findings become accountable actions, effects are verified and memory preserved.
4 Adaptive governance	Trends, scenarios and dependencies demonstrably change controls and strategy.
5 Resilient governance	Internal learning and external continuity are embedded in strategy, capital and oversight, with evidence over time.

Levels 4 and 5 require records of actual changes in decisions and controls. Policies, training, a certificate or a high average are insufficient. Organisational comparisons require disclosure of differences in evidence access and sensitivity analysis; incomplete records can systematically distort the composite.

4 Capability indicator manual

The following anchors constitute the consolidated v3.0 operational rules. For each indicator, record unit, period, cutoff, source IDs, E level, P or NE, confidence, RF and reasoning. Apply the anchor to the evidenced scope; do not extrapolate to the whole group.

GOV-01 Quality of governance challenge

Typical sources: Management and board minutes, alternative scenarios, dissent and decisions.

Rating	Operational anchor
P0	Affirmative operational evidence shows that material decisions are adopted without substantive challenge of alternatives and questions.
P1	Challenge is occasional, predominantly after an adverse event; follow-up of conclusions is inconsistent.
P2	Defined procedures require questions and alternatives; minutes demonstrate application with uneven depth.
P3	Independent challenge is traceable and changes assumptions, conditions, escalation or corrective activities.
P4	A time series shows that tested alternative scenarios influence strategy, capital and dependency mitigation.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

GOV-02 Escalation of material risks

Typical sources: Risk registers, escalation matrices and governing body materials.

Rating	Operational anchor
P0	Evidence shows that material risks remain outside oversight without a specified threshold, responsible person or deadline.
P1	Escalation depends on individuals and usually follows failure or external pressure.
P2	Thresholds, responsible persons and deadlines are defined; most material risks are recorded and escalated.
P3	Timeliness and exceptions are measured; governance oversight initiates documented action and follow-up of its effect.
P4	Weak signals and scenario thresholds are tested; data demonstrably improve preventive governance.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

GOV-03 Decision traceability

Typical sources: Decisions, approvals, supporting analyses and records of alternatives considered.

Rating	Operational anchor
P0	Material decisions cannot be reconstructed from the operational file because key authority, evidence or reasons are missing.
P1	Approvals exist, but evidence, reasons and assumptions are fragmented or depend on individual correspondence.
P2	Records identify authority, key evidence, assumptions and reasoning, with occasional gaps.
P3	Independent review can reconstruct alternatives, contrary evidence, conflicts of interest and subsequent activities.
P4	Historical reasons are preserved through changes and demonstrably used for learning, scenarios and improvements to later decisions.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

GOV-04 Oversight of external dependencies

Typical sources: Supplier, workforce, infrastructure and continuity maps.

Rating	Operational anchor
P0	Affirmative evidence shows that critical external dependencies are omitted from material governance oversight.
P1	Dependencies are considered reactively, after disruption or stakeholder pressure.
P2	Material dependencies are mapped, with responsible persons and basic continuity plans.
P3	Dependency trends, concentration and substitution tests regularly influence governing body decisions.
P4	Dependencies are embedded in strategy and capital allocation; mitigation adapts to actual or simulated disruption.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

CMP-01 Management of compliance obligations

Typical sources: Obligations register, regulatory updates and links to controls.

Rating	Operational anchor
P0	Evidence shows that material obligations are not reliably identified, assigned or maintained.
P1	Monitoring is informal or isolated by function; ownership, updating and links to controls are inconsistent.
P2	The register covers material obligations, responsible persons and controls; updating is carried out operationally.
P3	The register is ranked by risk; changes, exceptions and systemic gaps are regularly reviewed.
P4	Monitoring of emerging regulatory changes and risks leads to preventive changes in controls and resources.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

CMP-02 Response time to audit findings

Typical sources: Findings registers, management responses, deadlines and escalations.

Rating	Operational anchor
P0	Material findings have no approved response, responsible person or deadline; overdue findings remain uncontrolled.
P1	Responses are reactive and frequently late; formal closure takes precedence over effectiveness.
P2	Findings have approved plans, deadlines and responsible persons; implementation is monitored.
P3	Deadlines reflect risk; causes of delay are analysed, and verification of effect is planned.
P4	A time series shows preventive removal of bottlenecks and a reduction in high-risk overdue findings.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

CMP-03 Effectiveness of corrective actions

Typical sources: Retesting, action objectives, recurring findings and trends.

Rating	Operational anchor
P0	Affirmative evidence shows that actions are absent, nominal or materially non-functional.
P1	Activities are recorded, but completion is treated as success without reliable verification of effect.
P2	Objectives, responsible persons, deadlines and implementation evidence are defined; some actions are retested.
P3	Independent or objective testing checks recurrence and effects; ineffective actions are reopened.
P4	Longitudinal evidence shows sustained effect and transfers lessons into controls, incentives, resources or strategy.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

CMP-04 Investigation integrity

Typical sources: Mandate, investigation file, evidence protection, reasoning and oversight.

Rating	Operational anchor
P0	Operational evidence confirms a material lack of independence, evidence preservation or reasoned closure.
P1	A process exists, but mandate, scope, evidence handling, protection and consideration of contradictions are inconsistent.
P2	The protocol defines mandate, conflicts of interest, evidence retention, roles and closure reasoning.
P3	Independence, evidence inventory, competing explanations and links between evidence and reasons are auditable.
P4	Analysis across cases identifies systemic weaknesses and leads to demonstrable preventive protocol changes.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

SPK-01 Response to reports

Typical sources: Case management system, acknowledgements, triage, deadlines and feedback.

Rating	Operational anchor
P0	Operational evidence confirms that reports are not reliably received, assessed, assigned or processed.
P1	Channels exist, but responses depend on individuals; deadlines and status visibility are weak.
P2	Defined standards for acknowledgement, assessment, assignment and handling are implemented and monitored.
P3	Timeliness, case allocation and exceptions are analysed and independently reviewed.
P4	Capacity and trend analysis leads to sustained improvements in protection, escalation and response quality.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

SPK-02 Management of retaliation risk

Typical sources: HR records, monitoring of the reporting person’s position, surveys and independent reviews.

Rating	Operational anchor
P0	Evidence confirms the absence of systematic monitoring of adverse treatment or the disregard of credible signals.
P1	Protection is reactive and usually follows only an explicit allegation of retaliation.
P2	Defined controls, follow-up monitoring and selected HR and case indicators are in place.
P3	Signals are proactively monitored, independently reviewed and escalated; adverse actions have documented reasons.
P4	Longitudinal data demonstrate preventive protection and a reduction in indicators of fear or silence.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

SPK-03 Reporting person confidence

Typical sources: Segmented surveys, interviews, channel use and protection data.

Rating	Operational anchor
P0	Reliable evidence confirms that concerns cannot be raised safely or impartially. Unavailable measurements mean NE.
P1	Confidence is inferred from policy or report volume; fear, silence or perceived futility remain unresolved.
P2	Confidence is measured, and protection and roles are explained; results differ between groups.
P3	Segmented data are reviewed and linked to actions; fairness of handling and protection is monitored.
P4	Multiple independent sources over time confirm willingness to report difficult issues and a low level of fear.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

SPK-04 Learning at case closure

Typical sources: Closure reports, lessons registers and control changes.

Rating	Operational anchor
P0	Evidence shows case closure without assessment of systemic lessons, control weaknesses or recurrence risk.
P1	Lessons are recorded inconsistently; speed and the number of closed cases are the dominant criteria.
P2	Material cases include cause, action assessment and lessons; some lessons lead to control changes.
P3	Lessons are compared across cases, assigned to responsible persons and tested for effect.
P4	Learning is embedded in due diligence, training, controls, incentives and strategy with sustained recurrence monitoring.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

LRN-01 Preservation of institutional memory

Typical sources: Retention rules, linked repositories, access and searches.

Rating	Operational anchor
P0	Affirmative evidence shows that material warnings, reasons or action histories are not preserved or reconstructable.
P1	Knowledge depends on individuals, local files and informal memory.
P2	Defined repositories and rules preserve warnings, decisions, investigations and actions.
P3	Records are linked and searchable; completeness and continuity are checked through personnel changes.
P4	Memory survives changes in management or ownership and actively improves scenarios and decisions.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

LRN-02 Conversion of findings into learning

Typical sources: Improvement register, process changes and effect checks.

Rating	Operational anchor
P0	Operational evidence shows that material findings regularly remain without systemic improvement.
P1	Changes follow serious events, but remain local and are not monitored as conversion of findings into learning.
P2	Findings are assessed for the need for change; responsible persons and activities are recorded.
P3	Conversion rates, quality and effects are compared across functions; absence of change requires reasons.
P4	Evidence over time confirms embedded learning and reduced recurrence of governance and control weaknesses.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

LRN-03 Management of recurring failures

Typical sources: Incident taxonomy, earlier findings and responses before the information cutoff.

Rating	Operational anchor
P0	Materially similar failures recur without systemic recognition, linkage to cause or response.
P1	Recurrence is noticed after escalation; classification and trend analysis are weak.
P2	Recurrences are classified, monitored and linked to previous actions; thresholds trigger a response.
P3	Trends trigger independent review of causes, effectiveness, incentives and resource allocation.
P4	Evidence from the earlier frozen period confirms sustained improvements in recurrence management, controlling for exposure.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

Temporal boundary: LRN-03 codes earlier capability to recognise and address recurrence. The later recurrence rate is OUT-01. The same future data must not enter both baseline capability and the validation outcome.

LRN-04 Scenario adaptation

Typical sources: Scenario and continuity tests, resource changes and governance decisions.

Rating	Operational anchor
P0	Affirmative evidence shows that material new scenarios do not lead to governance action.
P1	Adaptation is occasional and begins after disruption becomes material.
P2	Defined exercises and scenarios inform selected governance activities.
P3	Scenario results change controls, resources, dependencies or thresholds before serious failure.
P4	Adaptive decisions are tested over time and embedded in strategy, using actual or simulated events.
NE	Evidence is unavailable or insufficient for the anchor. Zero is not assigned.

5 Exposure, controls and independent outcomes

Pair	ITE before mitigation	ECE tested controls
01	Supplier and customer concentration.	Diversification, substitution and continuity.
02	Dependency on critical inputs.	Input redundancy and tested continuity.
03	Local employment and community dependency.	Workforce transition and mitigation of local effects.
04	Infrastructure and essential services.	Redundancy, coordination and recovery.

ITE inherent exposure anchors

Rating	Inherent intensity
P0	Low or negligible structural dependency; practical substitution available.
P1	Limited dependency with several practical alternatives.
P2	Material dependency, partial concentration or switching barriers.
P3	High concentration, limited alternatives and significant local or infrastructure dependencies.
P4	Critical structural dependency with a credible pathway for rapid transmission of disruption.
NE	Insufficient evidence to characterise dependency.

ITE is determined before mitigation by specific organisational controls. Substitution availability as a market characteristic differs from a tested contract, backup capacity or organisational action, which belong to ECE. Sector adaptation must be documented in advance and must not change the scale direction.

ECE control effectiveness anchors

Rating	Verified effectiveness
P0	Affirmative evidence that relevant controls do not exist or are materially non-functional.
P1	Reactive or untested controls and informal improvisation.
P2	Defined controls are implemented; testing is partial.
P3	Independent tests and scenario exercises; weaknesses have verified corrective responses.
P4	Longitudinal evidence of adaptive controls, tested alternatives and sustained effect.
NE	Insufficient evidence of control effectiveness.

Outcome register

ID	Independent outcome after the baseline cutoff
OUT-01	Rate of materially similar recurring failures.
OUT-02	Time to a verified effective corrective action.
OUT-03	Disruption of production or service continuity.
OUT-04	Pressure on suppliers or creditors.
OUT-05	Employment instability or workforce transition outcome.
OUT-06	Financing or investment confidence outcome.

Legacy ECO-01–ECO-04 are retired as mixed constructs and separated into ITE/ECE pairs. ECO-05 is reclassified as OUT-06 and does not enter GRCS, ITE, ECE or RTE. Future outcomes are coded within a frozen window and must never be used to improve P4 retrospectively.

6 Empirical research and reliability

Research separately tests coding reproducibility, construct validity, longitudinal associations with outcomes and competing explanations. Assessor agreement does not confirm construct validity. Cross-sectional association does not establish causation.

Hypotheses and design

Proposed hypotheses link information traceability to institutional memory, reporting person protection to preservation of weak signals, conversion of findings into learning to lower recurrence, and GRCS to shorter time to verified effect. ITE may modify the relevance of capability weaknesses for continuity. GRCS, ITE and ECE must remain empirically distinguishable.

Before analysis, determine whether GRCS is a formative composite or a reflective latent construct. Tests must fit that decision. Address clustering of indicators within organisations, potentially non-random evidence missingness, detection intensity, duration of exposure, censoring of time to action, sector and macroeconomic shocks. Sample size requires a power or precision justification.

Sample	Plan in the source protocol	Separation
Development and calibration	4–6 practice and 6–12 development cases.	Excluded from reliability estimation and confirmatory validation.
Reliability	24–36 cases; two independent assessors.	Later follow-up only with frozen ratings and documented outcome blinding.
Internal validation	Power-based sample; planning ordinarily ≥100 organisations or cases.	Do not modify the manual after observing outcomes.
External replication	Separate team and institutions; justified power, planning ordinarily ≥100.	No overlap of organisations, assessors or adjudicators.

Two people independently code the same frozen package before reconciliation. Preserve original P, E, NE, confidence and RF ratings and reasons for disagreement. The case author or a party may supply evidence but should not be the sole confirmatory assessor. Where possible, keep outcome coders blinded to final GRCS/ITE/ECE results.

Agreement measurement plan

For ordinal P0–P4, use quadratically weighted Cohen's kappa alongside Gwet AC2, exact agreement, agreement within one category and disagreement of at least two categories. For binary NE and flag presence, use Gwet AC1, raw, positive and negative agreement, and Cohen's kappa. For domain composites, use absolute-agreement ICC(A,1). Cohen measures remain visible for comparability.

Level	Green retain	Amber revise
Ordinal indicator	N≥20; QWK≥0.70; lower 95% CI≥0.50; exact agreement≥0.50; within 1≥0.90; disagreement ≥2≤0.10; no systematic directional bias.	N≥12; QWK≥0.50; lower CI≥0.30; within 1≥0.80; no critical stopping reason.
Domain and composite	Complete N≥24; ICC≥0.75; lower CI≥0.60.	N≥18; ICC≥0.60; lower CI≥0.40.
NE classification	AC1≥0.80; agreement≥0.90; at least 5 records in each class.	AC1≥0.60; agreement≥0.80; rare class separately identified.
Red flag	AC1≥0.80; agreement≥0.95; positive agreement≥0.80; at least 5 positive cases.	Rare event or one threshold missed, without a material pattern of false-negative ratings.

Green requires all green conditions. Amber requires all amber minimums and no red condition. Otherwise, status is red and use is suspended or redesigned. Systematic confusion of NE and P0, a material pattern of false-negative flags and weight instability are stopping reasons. Fewer than five positive or negative events results in INSUFFICIENT_PREVALENCE where applicable; that status is not green.

Calculate confidence intervals by paired bootstrap at organisation level, with at least 2,000 replicates, a prespecified seed and a 95% percentile interval. The BCa interval is a sensitivity analysis. Record failed replicates and the effective sample count. Thresholds are research decision rules, not a universal definition of validity.

Outcomes and multiple testing

Group	Timing and analysis
Primary instrument	Independent duplicate baseline coding; QWK/AC2, ICC(A,1), NE and RF. No outcome interpretation.
Two co-primary outcomes	Recurrence and time to verified action, 12–24 months after baseline cutoff; one test family, Holm, two-sided α=0.05.
Secondary outcomes	Continuity, suppliers, employment and financing; frozen window; Benjamini–Hochberg FDR q=0.05.
Exploratory analyses	Sector interactions, RTE and alternative weights; clearly labelled effects and 95% CI.

Report effects, denominators, unadjusted and adjusted p/q values, and 95% intervals. Statistical significance is not practical validity. An unreported outcome cannot become primary after data inspection; additional analyses enter the deviation log.

Freezing and falsification

Freeze the protocol before real coding; anchors before calibration; the schema before the first import; the plan and code before outcome disclosure; outcomes and windows before outcome

access. Retain date, version and SHA-256, and a commit for code. Registration in a time-stamped public or controlled registry must precede access to actual outcomes. This publication does not claim that such registration has already occurred.

If assessors cannot distinguish NE from P0, revise the rules. If capability and exposure merge because of wording, revise the constructs. Narrow, split or retire persistently ambiguous indicators. If context better explains the outcome, retain the competing explanation. Report unsuccessful hypotheses; do not rescue them with post hoc weights.

7 Registers, evidence integration and implementation

Required registers

Register	Minimum fields
Sources	Source_ID; issuer/custodian; date; class; family; independence; provenance; authenticity; rights; reporting level; role; limitations; preservation.
Claims	Claim_ID; atomic claim; Source_ID; status; E; contrary evidence; competing explanation; evidence that changes the conclusion; bounded public wording; review.
Indicators	Unit; level; period; cutoff; ID; applicability; P/NE; E; confidence; RF; assessor; evidence; reasoning; exclusion reason.
Outcomes	Outcome_ID; date; post-cutoff window; measure and denominator; source; independence; prior risk pathway; causal limitation.
Samples	Organisation; one analytical role; team; calibration/validation/replication; overlap rule.
Versions	Version; change; reason; date; impact; review; preserved original; checksum; corrections channel.
Remedies	Case/finding; objective; responsible function; authority; deadline; implementation; retest; result; learning; post-transaction continuity.

Indicator definition sheet

For each indicator, complete stable ID and name, construct, observable definition, unit, period and cutoff, scale direction, sources, minimum E level, all P0–P4 anchors and the NE rule, specific RF, permitted sector adaptation and confidence rule. A change of meaning requires a new version; sector adaptation must not reverse direction retrospectively.

Accounting, fiscal and audit evidence

Methodological integration does not transfer specific ArcelorMittal claims into the core framework. K1_DOCUMENTED denotes a directly documented amount or event; K2_DERIVED denotes a reproducible derived value with preserved formula and assumptions; NV_UNVERIFIED denotes an unverified interpretation. These codes do not replace P, E or claim status.

Budget data: do not sum parent and subordinate classifications; remove extraction artefacts using a documented rule; separate budget plan, amendment, request, projection and execution. A fiscal outcome requires verified execution. A link to a particular company requires identified taxes, employment, transfers, suppliers or infrastructure.

Local reconciliations: preserve the general ledger, journal entry, date, account, AOP financial statement item mapping and the complete bridge between reporting layers. Do not infer

intent, authority, manipulation or loss from a difference without further evidence. Specifically request posting keys, audit logs, user rights, approvals, physical inventory records, final trial balance, tax reconciliation and filed statements.

A party’s expert report may contain reproducible arithmetic and an earlier risk signal, but the dual role of author and party must be recorded. Positive reported equity does not support a confirmed claim of negative total equity. An inventory difference across reporting constructs requires a roll-forward and classification bridge. A working capital movement in an indirect cash flow statement is not automatically cash from inventory sales. Debt-to-EBITDA requires reproducible EBITDA, debt composition and period data.

Group audit material: secondary analysis remains secondary until the original signed report, reporting entity and exact pages are archived. A group-level conclusion must not be transferred to local control effectiveness without a documented scope bridge. Hypotheses about deferred tax, intragroup financing or local harm require separate sources and a consistent reporting basis.

Field	Controlled value
REPORTING_LEVEL	PUBLIC_CONTEXT / GROUP / PARENT / LOCAL_ENTITY / TRANSACTION
DOCUMENT_ROLE	CONTEXT / ITE / ECE / OUTCOME / CLAIM / ASSURANCE_BOUNDARY
EVIDENCE_STATUS	K1_DOCUMENTED / K2_DERIVED / NV_UNVERIFIED
AUTHOR_INDEPENDENCE	INDEPENDENT / LIMITED / INTERESTED_PARTY / UNKNOWN
ASSURANCE_SCOPE	FINANCIAL_STATEMENTS / INTERNAL_CONTROLS / OTHER_INFORMATION / NONE
INFORMATION_CUTOFF	Immutable date preceding the validation outcome window.
PRIMARY_SOURCE_VERIFIED	YES / PARTIAL / NO
AGGREGATION_ELIGIBLE	YES only after hierarchy and non-overlap checks.
CAUSAL_STATUS	DESCRIPTIVE / ASSOCIATIONAL / CAUSAL_NOT_ESTABLISHED

Executable data controls

Check the unique organisation–indicator–assessor key and referential integrity. The same datum must not be both an input and a validation outcome. Reject OUT-06 in all formulas. The outcome date must follow the cutoff. Each organisation has one analytical role. Confirmatory analysis requires the corresponding protocol, anchor, schema and plan hashes.

Synthetic data carry SYNTHETIC_TEST_FIXTURE=YES, the SYN- prefix and a separate directory. Mixing actual and synthetic records stops confirmatory calculation. All synthetic charts, p values and results are labelled as software tests only. A rejected interpretation remains a traceable hypothesis or evidence request; it is not deleted from history.

Implementation cycle

The sequence is: lock boundaries and protocol; register sources; code independently; freeze agreement results; reconcile while preserving initial ratings; present separate constructs and coverage; assign actions and retest; code later outcomes; publish methods, limitations and corrections. Governing bodies should require a trail of actual implementation and effect; researchers should separate instrument reliability from validity and causation.

8 Release controls and sources

Control	Required record
G0	Frozen question, audience, unit, cutoff, risk and reporting level.
G1	Complete register of sources, provenance, families, independence and rights.
G2	Every material public claim linked to atomic evidence and the conclusion boundary.
G3	Contradictions, competing hypotheses, sequence, transition and falsification tests.
G4	Privacy, confidentiality, proportionality, reputational matters, opportunity to respond and editorial coherence.
G5	Named independent review, date, report, responses and resolution of material objections.
G6	Authorship and rights, frozen file, source and edition checksums, corrections channel and recorded release approval.

APPROVED means all applicable controls have passed. CONDITIONAL means external release awaits closure of open steps. REVISE requires changes to text or evidence; HOLD denotes an unresolved source, privacy, response, rights or contradiction issue; DO NOT RELEASE means the content cannot be responsibly substantiated. G5 is not synonymous with AI-tool review. This work does not declare G0–G4 complete without their registers.

Source lineage and editorial consolidation

The content is based on the supplied GRF v2.2, 76 pages; GRF v3.0 UNIFIED CONTROLLED PROTOCOL RC1 of 9 October 2026, 53 pages; FINAL AUTHOR CONTROLLED PUBLICATION EDITION of 12 October 2026, 32 pages; and its editable texts. Operational changes and their provenance are set out in separate document 02 of the Bosnian package. Repeated archival passages are not reprinted as competing rules. Source documents are retained separately.

This edition reconstructs operational tables, separate scales and registers; removes intermingled table sentences and duplicate rules. Historical version labels remain only in the lineage explanation. Citation and authorship references for this edition use v3.0. No independent reviewer report, research registration confirmation or release approval has been invented. This English translation preserves the scope, rules and status of the consolidated Bosnian edition.

Reference instruments

ISO 37000:2021, Governance of organizations — Guidance. Official description: https://www.iso.org/standard/65036.html. ISO 37301:2021, Compliance management systems — Requirements with guidance for use; and ISO 37002:2021, Whistleblowing management systems — Guidelines, are listed as reference instruments from the original GRF package. This edition does not reproduce their protected content or claim verification of compliance with individual clauses.

OECD (2023), OECD Guidelines for Multinational Enterprises on Responsible Business Conduct, OECD Publishing, Paris, https://doi.org/10.1787/81f92357-en. OECD (2018), OECD Due Diligence Guidance for Responsible Business Conduct, OECD Publishing, Paris, https://doi.org/10.1787/15f5f4b3-en. These instruments provide reference context; they do not confirm individual GRF ratings or the case.

Readiness record

Item	Status of this package
Editorial revision and version	Consolidated v3.0 text prepared in English.
Empirical validation	Not confirmed.
Independent review G5	Not documented in the available sources.
Authorship and legal clearance G6	Not documented in the available sources.
Case application	Separate; does not constitute instrument validation.
Corrections	Author’s channel provided in the supplement; record every substantive amendment in a new version.

Conclusions must remain bounded by evidence: a formal mechanism is not evidence of effectiveness, an evidence gap is not evidence of failure, and an adverse outcome is not evidence of causation. Resilience is examined through traceable governance activities that can be verified over time.