CHAPTER 20

QUALITY MANAGEMENT AND OPERATIONAL EXCELLENCE

Sixbert SANGWA

African Leadership University, Kigali, Rwanda | ssangwa@alueducation.com

PARENT BOOK
The Business Enterprise: Core Functions, Integrated Management, and Emerging Transformations
https://doi.org/10.65655/ocp.m30

VOLUME EDITOR AND PUBLISHER
Prof. Dr. Placide MUTABAZI
Open Christian Press, A publishing imprint of the Open Christian University and the Center for Faith and Work

CHAPTER DOI
10.65655/ocp.m30.c129

COPYRIGHT AND STATUS
© 2026 Sixbert SANGWA | CC BY 4.0

ABSTRACT

Quality is frequently reduced to inspection outcomes, certification status or isolated defect targets. None alone establishes that an enterprise reliably fulfils justified requirements without shifting avoidable risks to customers, workers and communities. This chapter develops a critical account of quality management and operational excellence across twelve interdependent domains, from competing quality philosophies and management systems to statistical variation, process diagnosis, lean, Six Sigma, quality economics, service recovery, supplier assurance, organizational culture and digital quality. A transparent, purposive synthesis combines seminal scholarship, peer-reviewed research, published international standards and Rwandan institutional evidence, while explicitly distinguishing conceptual claims from observed outcomes. The analysis incorporates the 2026 revisions of ISO 9000, ISO 9001, ISO 14001 and ISO 19011, the current transition arrangements and limitations of certification. A constructed Kigali packaging case supports reproducible Pareto, p-chart, cost-of-quality, investment, risk and corrective-action exercises. An original accountability architecture links credible specifications, capable work, measurement validity, employee voice and verified learning. The Rwanda–Toyota comparison identifies the complementary resources and institutional conditions that limit simple transfer of management prescriptions. Quality is consequently defended as a technically informed, morally constrained and empirically revisable enterprise capability rather than an administrative label. The proposed synthesis is analytical and educational, not an empirically validated causal model.

Keywords: quality management; operational excellence; statistical process control; root cause analysis; lean management; Six Sigma; quality costs; service quality; supplier assurance; Rwanda.

LEARNING OUTCOMES AND CONCEPTUAL ROADMAP

After completing this chapter, the reader should be able to (LO1) distinguish conformance, fitness for use, dependability, perceived value and legally relevant safety; (LO2) critically compare inspection, assurance, TQM, lean and Six Sigma and specify their assumptions and limitations; (LO3) develop a proportionate quality management system with process ownership, documented controls and audit evidence; (LO4) calculate and interpret defect frequencies, Pareto shares, control limits and qualified capability indices; (LO5) design a testable corrective-action procedure that distinguishes symptoms from established causes;

(LO6) estimate prevention, appraisal and failure costs and evaluate a quality investment under uncertainty; (LO7) evaluate service, supplier, digital and environmental quality against fairness, security and institutional obligations; and (LO8) defend a cross-functional improvement proposal for a Rwandan enterprise against credible international comparators.

The analysis follows the approved sequence, beginning with the contested meanings and intellectual history of quality (20.1–20.2), advancing through management systems, process evidence and diagnostic methods (20.3–20.5), examining improvement traditions and economics (20.6–20.8), and extending the analysis to customers, suppliers, people and technology (20.9–20.12). Sections 20.13–20.16 integrate an auditable Rwanda teaching case, international comparison, critical synthesis, research agenda, terminology and assessment resources. Chapter 17 addressed operating capacity and process flow; Chapter 18 covered source-to-contract governance; Chapter 19 developed inventory and logistics. The present chapter owns assurance, defect prevention, variation analysis and corrective action, not academic quality assurance or specialist health-sector clinical regulation. Chapter 21 will address temporary project delivery rather than recurring quality routines.

20.0 INTRODUCTION: WHEN AN ACCEPTED ORDER IS NOT AN ACCEPTABLE PRODUCT

Consider a hypothetical food-packaging enterprise supplying shops and institutional buyers in Kigali. Production reports 12,000 packets completed in a month, logistics reports dispatch on schedule, and sales reports that customers have reordered. Yet several distributors have received packets with weak seals; internal checks have repeatedly found incorrect labels and underfilled packs. The enterprise may meet an output target while transferring inspection, replacement and reputational costs downstream. It may also respond by demanding more final inspection, thereby detecting failure later without changing the conditions that generate it. The enterprise is invented for instruction; neither its results nor its defects describe an identified business.

Quality management concerns the governed processes by which an organization establishes requirements, designs the means of meeting them, verifies evidence of conformity and fitness, responds to failures and revises its system. Operational excellence is a wider and more contestable aspiration: the repeatable ability to deliver legitimate value reliably while improving flow, safety, cost, adaptability and accountability. This distinction matters. An apparently efficient process may consistently produce an undesirable item; a technically conforming product may disappoint an unmet user need; and a satisfied customer may not detect a hidden safety hazard. Quality cannot be inferred from any single point in this chain (Garvin, 1987; Juran, 1988; International Organization for Standardization [ISO], 2026a).

The chapter advances a conditional proposition: superior quality performance is produced neither by inspection volume nor a management certificate in isolation, but by the fit among justified requirements, capable processes, sound measurement, effective organizational challenge and remedies proportionate to the harm at stake. This view is deliberately broader than quality as cost minimization. It recognizes non-negotiable constraints where an average yield or favorable expected return cannot legitimate a hazardous product, false test report, suppressed complaint or unjust transfer of risk. The practical burden is to make different claims explicit: what was specified; what was measured; what was inferred about causes; what was decided; and who remains responsible if the inference is wrong.

20.0.1 Evidence strategy and limits of claims

This chapter uses purposive critical synthesis of secondary sources, not an experimental study, enterprise survey or preregistered systematic review. The selection logic combined original texts that created competing quality traditions, peer-reviewed studies on relationships between practices and outcomes,

current source institutions for standards and professional procedures, and Rwandan authorities for country-specific context. Targeted author-title, DOI and institutional-source checks undertaken on 2 November 2026 covered Deming, Shewhart, Juran, Flynn et al., Powell, Samson and Terziovski, Kaynak, Shah and Ward, Zu et al., Corbett et al., ISO, the American Society for Quality (ASQ), the National Institute of Standards and Technology (NIST), Rwanda Standards Board (RSB), and the National Institute of Statistics of Rwanda (NISR). This is a reasoned selection, not exhaustive retrieval; it does not assert PRISMA screening numbers, meta-analytic effect sizes or new primary observations. The numerical enterprise records below are expressly constructed to teach reproducible methods, while institutional statistics retain their stated coverage and reference years.

20.0.2 The philosophical and inferential problem of quality

Quality judgments involve an account of the good being pursued, an account of what can be known about an output, and a defensible way to adjudicate competing claims. The first issue is teleological: a process can conform perfectly to a specification and nevertheless be organized to deliver an unsuitable or harmful product. The second is epistemic: measured agreement with a specification establishes only what the measurement system can validly detect under its sampling and decision rules. The third is distributive: firms can report savings while households, subcontractors or ecosystems bear costs that their internal quality accounts omit. Consequently, specification, verification and moral justification must not be collapsed into one managerial score. Donabedian's (1988) distinction among structure, process and outcome, developed for healthcare, is instructive as an analytical comparison rather than a template to be imposed uncritically on manufacturing: good structures are enabling conditions, observable processes are intermediate mechanisms, and experienced outcomes are the ultimate objects requiring independent inquiry.

This philosophical discipline also guards against the substitution of management proxies for purposes. A certificate attests to a defined management system within an assessed scope; a control chart offers conditional evidence about process behavior; a customer rating reports an experience filtered through expectations, access and response patterns. None grants its holder epistemic authority over all relevant dimensions of quality. The chapter therefore treats quality management as a practice of fallible but accountable judgment: propositions are useful when they can be contradicted by credible evidence, and decisions remain justified only while their requirements, measurements and consequences withstand scrutiny. In African business environments, this entails recognizing tacit production knowledge and informal exchange while refusing to romanticize arrangements that lack effective protection, traceability or remedy.

20.1 THE MEANING OF QUALITY: REQUIREMENTS, EXPERIENCE AND RESPONSIBILITY

20.1.1 Competing concepts and their practical consequences

Quality is a relational, multidimensional judgment, not an intrinsic label that a product or service possesses without reference to purposes and requirements. Conformance means that the observed output satisfies specified characteristics within permitted tolerances. Fitness for use asks whether those characteristics actually enable users to achieve their purposes. Customer-perceived quality includes interpreted performance and experience, but is influenced by expectations and information asymmetry. Reliability concerns performance over time under stated use conditions; safety concerns the acceptability of hazards and their controls. These dimensions can disagree. A packet with the correct weight may have an unsafe seal; a safe service can still be unacceptably late; a luxurious product can be well liked while concealing exploitative or environmentally damaging production.

Juran's managerial emphasis on fitness for use and Crosby's conformance-oriented prevention argument illuminate distinct decisions rather than forming a simple contest over the single correct definition (Juran, 1988; Crosby, 1979). Garvin (1987) identified product-quality dimensions including performance, features, reliability, conformance, durability, serviceability, aesthetics and perceived quality. Their relative weight differs with product, customer group and regulatory context. For a medicine or a food product, a pleasant appearance cannot compensate for an unacceptable safety risk. For a public-facing administrative service, correct substantive results may coexist with inaccessible procedures or unfair treatment. These differences establish the necessity of explicit requirements and a defensible priority structure.

20.1.2 From a customer's voice to an auditable requirement

A useful quality specification identifies the characteristic, measurement rule, acceptance limit, intended use, responsible decision maker, applicable legal or contractual source, and evidence to be retained. An order stating 'adequately sealed' is not a controlled acceptance criterion until seal integrity has an operational test, calibrated method and decision rule. Likewise, 'prompt customer service' needs an agreed time window, defined start and stop events and explicit handling of exceptional cases. Requirements should emerge from customer needs, engineering feasibility, public obligations, worker knowledge and foreseeable misuse, not only from the preferences of a major account.

Quality function deployment and related voice-of-customer methods can connect expressed needs to technical design, but conversion is not mechanically neutral. Interview samples may omit persons with disabilities, smaller distributors or informal customers; employees closest to a hazard may have little formal authority; and customer demand can conflict with safety. Responsible design therefore distinguishes desiderata that may be traded against cost from mandatory controls whose violation is unacceptable. Measurement and governance must also distinguish the object being counted: a defective unit, a defect opportunity, a complaint, a customer account and an escaped hazard are not interchangeable denominators.

20.1.3 Quality and operational excellence are not synonymous

Operational excellence encompasses flow performance, adaptability, asset utilization, customer trust and resource stewardship. Quality management contributes to these outcomes, but may impose necessary inspection time, redundancy or segregation that appears wasteful under a narrow speed metric. An independent release decision may delay shipment precisely because doing so protects a higher-value objective. This becomes a strategic question about which outcomes are optimized and which constraints are inviolable, consistent with Chapter 6's emphasis on coherent commitments and Chapter 8's concern for accountable decision rights. Excellence is a claim requiring evidence across dimensions, not a prize conferred by a slogan or a temporary reduction in reported errors.

20.2 THE EVOLUTION OF QUALITY MANAGEMENT: FROM SORTING TO SYSTEM LEARNING

20.2.1 Inspection, statistical control and managerial responsibility

Mass production expanded the challenge of detecting nonconformity after work had been separated into specialized tasks. End-of-line inspection can segregate defective units, but does not itself establish which upstream mechanism created them. Walter A. Shewhart's control-chart work in the 1920s and his 1931 book shifted attention from isolated nonconforming outputs toward variation in the generating process (Shewhart, 1931; ASQ, n.d.-a). His intellectual contribution was not a guarantee of perfect prediction: it was a way to distinguish patterns reasonably compatible with a stable process from evidence suggesting an

assignable disturbance. It thereby transformed diagnosis into a sequential decision about when to intervene.

W. Edwards Deming carried statistical thinking into an explicitly managerial account of systems, learning and leadership. His criticism of numerical targets divorced from methods was directed at organizations that blamed individual workers for variation produced by their work systems. Deming argued that improvement requires managerial redesign of interdependent processes, not simply stronger inspection or worker exhortation (Deming, 1986). Joseph Juran gave equal importance to planning quality, controlling performance and deliberately undertaking improvement, emphasizing managerial responsibility for fitness for use. Their approaches diverged in terminology and emphasis but both rejected the idea that quality resides solely within the inspection department (Juran, 1988).

20.2.2 Organization-wide quality and empirical uncertainty

Later total quality management (TQM) articulated a cross-functional approach involving leadership, employee participation, customers, supplier relationships, process management and continuous improvement. The attraction was its systemic ambition; its weakness was inconsistent operational definition. Flynn et al. (1994) developed a measurement framework partly because broad claims about 'having TQM' lacked comparable constructs. Powell (1995) argued that certain socially complex practices, particularly commitment and culture, might contribute to advantage in ways that readily imitated tools did not. Samson and Terziovski (1999), Kaynak (2003) and Sila and Ebrahimpour (2005) reported relationships between quality-management configurations and performance, but many designs relied on organizational surveys, measured interrelated practices simultaneously, and could not establish universal intervention effects or eliminate selection and reverse causation.

These studies support a careful conclusion rather than a managerial guarantee. Quality practices are more plausibly effective as complementary systems than as isolated checklists, yet observed associations need not imply that importing the vocabulary, purchasing a certificate or copying a high-performing firm will create equal results. Organizational history, capital, product risk, leadership credibility and data quality matter. A tiny Rwandan processor cannot assume the same specialist staffing and instrument costs as a multinational plant. Nor does informality imply an absence of skill: some microenterprises possess sophisticated tacit inspection routines while lacking written calibration, traceability and independent assurance.

20.2.3 Assurance, certification and standards in transition

Quality assurance shifts emphasis to planned confidence in a system's capacity to satisfy requirements, while quality control checks whether specified results do satisfy them. ISO 9000:2026 provides quality-management fundamentals and vocabulary, whereas ISO 9001:2026 specifies requirements for certifiable quality management systems; the two have distinct purposes (ISO, 2026a, 2026b). ISO lists ISO 9001:2026 as published in September 2026, superseding its 2015 edition, and identifies transitional arrangements to be handled through certification bodies. At the time of this chapter, some Rwanda institutional pages still discuss RS ISO 9001:2015; such pages cannot on their own establish the status or transitional obligations of a particular organization's certificate (ISO, 2026b; RSB, n.d.-a). ISO 19011:2026 provides guidelines for auditing management systems and is not a product-quality certificate (ISO, 2026c).

Certification is neither synonymous with legal compliance nor direct evidence that every individual output is safe. The certification scope, issuing body's competence and accreditation, surveillance arrangements, applicable edition, and condition of real processes matter. Corbett et al. (2005) found financial improvements following ISO 9000 certification among matched U.S. manufacturing firms, yet the

magnitude was sensitive to comparison specifications. It would be unsound to treat this as proof of the same causal return for a small African service provider or to infer that certification caused all subsequent improvement. For sector-specific hazards, additional systems may be required: ISO 22000:2018 concerns food safety management and integrates hazard-analysis principles, but cannot be treated as a substitute for every applicable product regulation or hygiene requirement (ISO, 2018a).

20.2.4 Quality philosophies as rival accounts of change

Although they share a commitment to reducing avoidable failure, major quality traditions make different assumptions about the locus of improvement. Statistical process control locates much of the immediate evidentiary problem in temporal variation and rational subgrouping. Juran's planning–control–improvement distinction highlights management's obligation to build quality into purposes and processes. Crosby's prevention-oriented conformance approach can clarify standards but risks narrowness if the specification itself is inadequate. TQM requires organizational complementarity and credible worker involvement; lean emphasizes flow and the design of work; Six Sigma formalizes analytical project discipline. Their claims should therefore be compared at the level of mechanisms rather than by popularity or brand. Hackman and Wageman (1995) warned that managerial rhetoric often aggregates unlike activities under the TQM label without evidence that the same intervention works in different contexts.

A further distinction separates improvement through learning from improvement through authority. Managers can compel reporting, increase inspection frequency and purchase instruments; they cannot compel a valid finding or genuine candour merely by issuing a policy. A credible quality programme has to distribute rights to observe, challenge, stop, repair and learn. Such rights are economically important because local knowledge about process variation often resides with people who do not control capital expenditures. The implication is not that participation always substitutes for technical expertise; metrology, statistical design and safety judgment require competence. Rather, expert verification and frontline experience provide partially independent forms of evidence whose productive combination needs deliberate organizational design (Deming, 1986; Hackman & Wageman, 1995).

20.3 QUALITY MANAGEMENT SYSTEMS: OPERATING ARCHITECTURE AND ASSURANCE

20.3.1 The process-based system

A quality management system (QMS) is the interconnected collection of policies, responsibilities, processes, records, controls and review mechanisms used to achieve the organization's quality objectives. An effective design starts with external and internal context, relevant interested parties, product and service requirements and the risks of failure. Its scope must match actual processes, locations and offerings rather than the most attractive operations to present during audit. Process maps should specify input conditions, activities, decision gates, permissible variation, handoffs, outputs, owners, controls and records. Documented information should exist because it protects memory, traceability or accountability, not because a consultancy sells document volume.

Leadership establishes policies, objectives, accountability and resources. Process owners demonstrate control in everyday decisions; competent employees maintain and challenge working methods; assurance personnel test whether recorded practice and actual practice agree. For high-consequence releases, a person with appropriate technical authority must be able to stop shipment despite sales pressure. This allocation is a governance choice. In a microenterprise the founder may hold several roles, but incompatible acts such as performing a disputed test and independently approving its validity should be separated where risk warrants. Where separation is financially infeasible, an external test, peer check or documented independent review can provide proportionate assurance.

20.3.2 Control of documented information and change

A lean QMS needs an accessible requirements register, document-version control, approved operating instructions where consequential, product identification and traceability, calibration evidence, competence records, nonconformity and corrective-action registers, complaint records, internal audits, management reviews and change approvals. Each document should identify owner, status, applicable scope and retention logic. Staff need practical access to the current approved version at the work point. A procedure saved on a manager's private computer, or written only in a language its intended users cannot read, cannot reliably control work.

Changes in packaging material, suppliers, recipes, software models, customer requirements or operating shifts should trigger risk review before release. Change control asks what requirement may be affected, what verification is necessary, how stock produced under the old design is distinguished, and who approves the transition. An emergency deviation may sometimes be justified, but it must be authorized, time-limited and traceable. The number of completed forms is not a suitable proxy for this system's effectiveness: meaningful evidence includes reduced process instability, detected near misses before release, timely corrective actions, and whether employees can demonstrate the procedure without coaching.

20.3.3 Audit design and assurance limits

Internal audits should follow risk rather than identical annual coverage of every low- and high-consequence process. A risk-based audit programme defines criteria, scope, auditors' competence, independence where practicable, sampling rationale, records and follow-up. Evidence should triangulate documents, observation, interviews and transaction traces. One signed checklist may demonstrate that a form exists, not that the process produces safe outputs. ISO 19011:2026 supplies audit guidance, while external accredited certification operates under separate conformity-assessment rules (ISO, 2026c; RSB, n.d.-b). Management reviews should consider complaints, process trends, supplier performance, audits, corrective-action effectiveness, resource adequacy and emerging changes, and record decisions with owners and due dates.

The major failure mode is ritual compliance. A manager can obtain favorable audit scores by selecting convenient samples, closing cases at the 'retrain employees' stage, or concealing defect data. Consequently, a defensible QMS must protect reporting channels and look for contradictions: rising scrap combined with mysteriously declining recorded defects; perfect test results immediately after a supplier change; or no complaints where customers lack an accessible complaint channel. The most credible audit is a disciplined challenge to organizational assurance, not a rehearsal for a certificate photograph.

20.3.4 Risk-based objectives, management review and certification transition

A mature QMS connects objectives with decision thresholds. The organization should decide which characteristics are critical to safety, which drive customer use and which concern discretionary preferences; establish verification depth proportionate to consequences; and document who can accept, reject or escalate an exception. Management review must then interpret changes in complaints, rework, escapes, supplier lots, corrective actions and resource constraints against their measurement quality. A favorable aggregate defect rate cannot compensate for an uncontrolled critical hazard, while a deterioration in reporting counts may reflect a healthier detection culture rather than an underlying increase in failure. Audit agendas should deliberately include contrary evidence, such as products rejected despite strong delivery indicators or employee reports that contradict management summaries.

The September 2026 publication of ISO 9001:2026 introduces a practical transition rather than instant invalidation of all prior certificates. Published ISO transition guidance indicates that organizations already certified to ISO 9001:2015 have until 30 September 2029 to complete transition, while new and initial accredited certificates are to be issued only against the 2026 edition from 31 March 2028; related accreditation and certification-body deadlines precede these dates (ISO/TC 176, 2026). The validity of any individual certificate nevertheless depends on its issuer, accreditation, scope, expiry, surveillance and transition circumstances. A Rwandan enterprise should verify its actual certification body's instructions and nationally adopted standard rather than infer compliance merely because an older web page remains online. These date-sensitive rules are reported as of 2 November 2026, not as timeless requirements.

20.4 PROCESS VARIATION, MEASUREMENT AND STATISTICAL INFERENCE

20.4.1 Stable and unstable processes

Every repeated process exhibits variation. Common-cause variation represents the routine, jointly produced variability of a comparatively stable system; special-cause variation is an unusual shift or disturbance that warrants specific inquiry. The distinction is statistical and operational, not an assertion that ordinary variation is ethically acceptable. A stable process may consistently produce too many unsafe units; statistical control establishes predictability under assumptions, not conformity with customer limits. Conversely, an isolated outlier may result from a measurement mistake, a supplier lot, a staff substitution or a genuine process failure. Shewhart's logic warns against tampering, whereby managers adjust a stable process in response to ordinary fluctuations and unintentionally increase variability (Shewhart, 1931; NIST/SEMATECH, n.d.).

Control charts arrange time-ordered measurements against a centre line and limits derived from an assumed in-control process. Attribute charts include p-charts for fraction of inspected units nonconforming and u-charts for defects per unit when opportunity exposure is appropriately represented. Variable charts follow measurements such as seal strength or weight, with chart selection determined by subgroup size, distribution and sampling design. A run chart may be an honest first instrument where the history is insufficient to estimate limits. It is methodologically incorrect to derive a 'three-sigma limit' from a few arbitrary observations, to silently remove inconvenient observations, or to compare a statistical control limit with an engineering specification limit as if they served the same decision.

20.4.2 Worked construction of a p-chart

For m subgroups with n_i inspected units and d_i nonconforming units, the pooled estimated fraction nonconforming is $\bar{p} = \Sigma d_i / \Sigma n_i$. When independent Bernoulli observations within a rational subgroup are a reasonable approximation, the subgroup chart uses $UCL_i = \bar{p} + 3\sqrt{[\bar{p}(1 - \bar{p})/n_i]}$ and $LCL_i = \max\{0, \bar{p} - 3\sqrt{[\bar{p}(1 - \bar{p})/n_i]}\}$. NIST's Engineering Statistics Handbook gives the underlying binomial logic (NIST/SEMATECH, n.d.). Clustering, overdispersion, short runs, changing inspection intensity and rare events can invalidate naïve limits; technical judgment must precede interpretation.

The constructed teaching case later presents eleven baseline samples of n = 200 each with 91 nonconforming units, so $\bar{p} = 91/2{,}200 = 0.04136$. The estimated three-sigma control limits are LCL = 0 and UCL = 0.08361. A twelfth sample contains 23 nonconforming units among 200, or 0.115, which exceeds the baseline UCL. The reasonable conclusion is that the later subgroup is not adequately explained by the baseline model, not that a particular operator caused the deterioration. A defensible investigation examines shift, material lot, inspection procedure, machine settings and workload before declaring a cause or recalculating baseline limits.

20.4.3 Measurement-system validity and capability

Measurement systems have error arising from instruments, calibration, environmental conditions, sampling, operator judgments, definitions and data entry. Gauge repeatability and reproducibility studies help separate instrument and appraiser variation, while attribute agreement studies are relevant where inspectors classify defects. If two inspectors disagree systematically about a 'weak seal', adding more data will create false precision. Before comparing suppliers or claiming improvement, management should test whether the measurement method detects the relevant range of defects consistently and independently of who benefits from the reported result.

For approximately normal, stable measurement distributions, capability indices may be expressed as $C_p = (USL - LSL)/(6\sigma_{within})$ and $C_{pk} = \min[(USL - \mu)/(3\sigma_{within}), (\mu - LSL)/(3\sigma_{within})]$, where μ is process mean, σ_{within} is a defensible estimate of within-process standard deviation, and LSL and USL are engineering specification limits. These indices do not themselves establish process stability, normality, adequate measurement or safety; a high index for a poorly chosen specification is of little public value. For skewed data, autocorrelation or mixtures of production modes, distribution-free or appropriate alternative capability analyses should be evaluated. An enterprise should never announce 'six-sigma quality' merely because an index exceeds a target without an explicit model for defects and opportunities.

20.4.4 Sampling, decision risks and the difference between detection and assurance

Acceptance sampling addresses whether a defined lot should be accepted under a specified sampling rule; process control addresses whether the generating process has changed. These are different inferential decisions. Consider a hypothetical lot containing 5% nonconforming units, with a simple random sample of 20 treated as independent draws for an approximate binomial calculation. A rule that accepts the lot only when no defective unit is found would accept such a lot with probability $(0.95)^{20} = 0.3585$, or about 35.8%. The numerical example illustrates why passing inspection does not establish zero defects, and why a superficially strict zero-acceptance rule can still miss substantial problems when the sample is small. Exact finite-lot inference would instead use the hypergeometric distribution, because sampling without replacement alters the probability. Neither model is valid if the sample is taken from convenient locations or if inspection sensitivity is imperfect (Montgomery, 2020).

In practice, sampling plans should distinguish the producer's risk of rejecting an acceptable lot from the consumer's risk of accepting an unacceptable one, using an operating-characteristic curve with quality levels and decision rules appropriate to the actual risk. A single acceptance-quality-limit label is not a guarantee that a lot has that quality; nor may statistical sampling replace mandatory testing of safety-critical characteristics. In settings where consequences are severe, containment, hazard analysis and traceability may require different release controls. When the parameter is a continuous characteristic, the unit's measurement uncertainty and its position near a legal or engineering limit must be considered, not merely whether a displayed number falls inside the tolerance.

20.4.5 Reproducible capability, precision and uncertainty

Two distinctions improve the integrity of quality statistics. First, process capability must be evaluated after stability and measurement suitability have been established; a computed C_{pk} from unstable or mixed processes may conceal systematic change. Second, precision of an estimated defect proportion is not the same as adequacy of its sampling design. Under independent sampling, a rough standard error for the pooled baseline proportion in the teaching case is $\sqrt{[0.04136 \times 0.95864 / 2{,}200]}$, about 0.00425; that quantity describes sampling variability around the pooled estimate, not the between-subgroup variation addressed by the p-chart. Using it interchangeably with three-sigma subgroup limits would understate the

practical distinction between estimation and monitoring. Serial dependence, within-shift clustering, imperfect classification and discretionary sampling can defeat both calculations.

A transparent quality-data register should therefore identify the sampling frame, rational subgroup definition, sample size, instrument, calibration state, inspector, classification rule, missing observations, exclusions and reasons for any change in measurement. Before-and-after analyses also need an explicit statement of whether product mix, inspection intensity or customer exposure changed. These details determine what the statistic means. They are especially important when quality systems migrate from handwritten check sheets to digital sensors, because faster data collection can increase apparent defect counts solely by detecting events previously unobserved. It would be analytically mistaken to describe that increase as deterioration without a measurement bridge (NIST/SEMATECH, n.d.; Montgomery, 2020).

20.5 ROOT CAUSE ANALYSIS: FROM OBSERVATION TO CORRECTIVE ACTION

20.5.1 Diagnostic logic and the seven basic tools

A root cause is not simply the last answer produced by asking 'why' five times. A meaningful causal account identifies a factor or set of mechanisms for which a feasible intervention would reduce recurrence under specified operating conditions. The seven basic quality tools commonly include a check sheet, histogram, Pareto chart, cause-and-effect diagram, scatter diagram, control chart and stratification or a closely related process chart (ASQ, n.d.-b). Their legitimate purpose is to structure evidence and generate testable hypotheses. A Pareto chart ranks counted categories; it does not establish that the largest count is the easiest, safest or most economically justified cause to address. An Ishikawa fishbone organizes hypotheses, not proof.

A useful failure statement identifies the affected characteristic, exact operational definition, location, time period, denominator, severity and existing detection point. Investigators then trace the process, observe work at the point where nonconformity arises, stratify by machine, shift, input lot, operator or route, and compare candidate explanations with disconfirming observations. Correlations between lateness and a supplier do not establish that supplier inputs caused underfill if both changed during a production surge. A 'five whys' exercise is strongest when each link is checked against documents, physical observation or a controlled comparison. Multiple interacting causes are ordinary; a single linear chain can misrepresent a complex work system.

20.5.2 Correction, corrective action and effectiveness verification

Correction deals with an identified nonconforming output: segregating, relabelling, reworking, recalling or refunding it as appropriate. Corrective action changes the cause-producing system: modifying a sealing temperature control, qualifying a new supplier specification, redesigning the inspection method or clarifying a contradictory operating instruction. Preventive thinking extends beyond the immediate failure to analogous processes, but improvement teams should not substitute speculative risk lists for proven hazard controls. The responsible process protects customers first where safety is uncertain, preserves traceability and documents who authorizes product disposition.

Effectiveness review is the step most vulnerable to performative closure. Completion of a training session does not prove that sealing defects declined; closure requires a predefined outcome, exposure denominator, observation period, performance criterion and monitoring responsibility. A pilot may use the same chart metric before and after the change while also recording possible shifts in product mix and inspection intensity. If the intervention coincides with changes in demand or equipment, causal attribution

remains provisional. An appropriate response is to document the uncertainty, maintain protective controls and revise the hypothesis, not to present an attractive before–after chart as a randomized experiment.

Figure 20.1

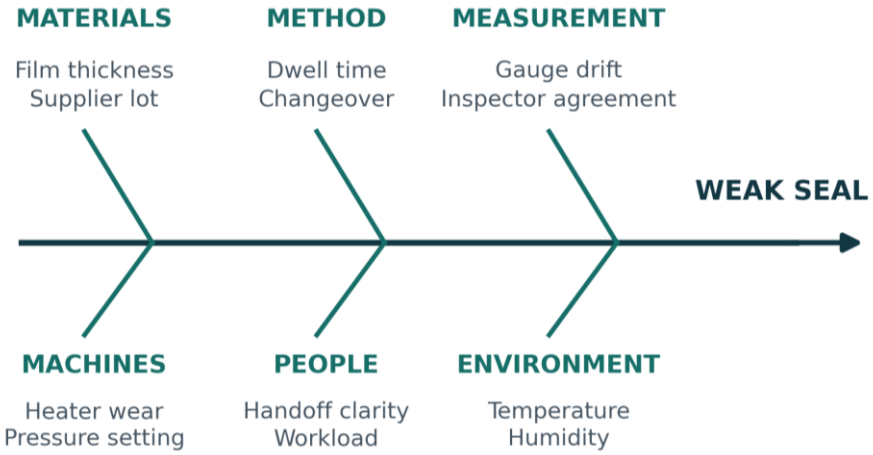
An accountable plan-do-check-act (PDCA) learning loop



Note. Original explanatory figure synthesizing the cited concepts; all case values, where shown, are hypothetical.

Figure 20.2

Ishikawa hypothesis map for weak package seals



Note. Original explanatory figure synthesizing the cited concepts; all case values, where shown, are hypothetical.

20.5.3 Risk-based failure analysis without false numerical certainty

Failure mode and effects analysis (FMEA) can supplement direct causal investigation by systematically considering the function that might fail, its credible modes of failure, consequences, initiating mechanisms, existing prevention and detection controls, and feasible corrective actions. For the hypothetical sealing line, loss of temperature control, unsuitable film, poor seal-surface contact and incorrect inspection classification should be documented as distinct candidates. Their consequences may range from cosmetic appearance to product contamination and customer harm. The investigation must ask how each failure could escape detection and what would interrupt the chain before release. This prospective analysis complements the retrospective examination of observed seal failures; it cannot identify an actual cause merely because a workshop has assigned one a high score.

Traditional FMEA templates sometimes multiply ordinal scores for severity, occurrence and detectability into a risk-priority number. Multiplication assumes a level of numerical comparability that ordinal rankings usually do not possess: 8 × 3 × 3 and 3 × 8 × 3 may yield identical totals despite fundamentally different safety implications. Particularly severe failure modes need explicit action criteria regardless of aggregate

rank, and estimates should disclose uncertainty where historical evidence is sparse. The actionable output is a prioritized set of technically justified safeguards, named owners, verification evidence and escalation thresholds, not a mathematically impressive table that obscures the distribution of harm. Work instructions, staffing, supplier inputs and product design should all remain possible intervention targets (Montgomery, 2020).

20.6 LEAN PRINCIPLES AND WASTE REDUCTION

20.6.1 Flow, value and the limits of 'waste'

Lean practice seeks to organize work so that legitimate customer value can flow with fewer unnecessary delays, corrections, movements, inventories and interruptions. Associated practices include observing the value stream, standard work, visual control, workplace organization, pull replenishment, equipment reliability and structured problem-solving. The Toyota Production System's documented pillars are just-in-time and jidoka, the ability to bring human judgment and stopping capability into automated production (Toyota Motor Corporation, n.d.). Lean should not be interpreted as making every employee continuously busy or eliminating all buffers: queueing variability and health, safety or supply risks can make controlled reserves indispensable. Shah and Ward (2007) demonstrated that lean is better represented as a configuration of mutually reinforcing practices than as a single intervention.

The familiar waste taxonomy includes defects, overproduction, waiting, unnecessary transport, excess processing, inventory, motion and underused human knowledge. However, waste requires a stated viewpoint. A duplicate check may be unnecessary for low-risk stationery but necessary for an irreversible clinical or financial decision. Slack capacity may allow safe recovery after equipment stoppage. A culture that treats worker rest, reasonable accessibility adjustments or paid learning as 'non-value-adding' can improve a local throughput ratio while degrading long-term reliability and justice. Lean analysis should distinguish resources that are truly squandered from protective capacity, ethical safeguards and work investments with delayed returns.

20.6.2 Practical deployment in resource-constrained enterprises

A modest enterprise may begin with one value stream, visible work standards, labeled defect bins, daily short reviews and direct worker observations rather than expensive software. The most consequential prerequisites are a stable operational definition of success, accurate signals, psychological permission to stop unsafe work and the ability to correct recurring obstacles. Standard work documents the current best-known safe method while preserving an explicit process for frontline suggestions. Value-stream mapping can expose repeated waiting between packaging and release, but estimates must be tested against observed timestamps; a visually persuasive map is not automatically a valid time study.

Lean implementation can fail through aggressive headcount cutting, perfunctory 'kaizen events', shifting stock to suppliers or bypassing inspection to meet daily targets. Toyota's practices emerged through a long history of supplier relationships, technical routines and organizational learning; transferring isolated artifacts to Rwanda without matching skills, spare-part availability, maintenance and supplier conditions may produce imitation rather than effectiveness (Toyota Motor Corporation, n.d.; Shah & Ward, 2007). The correct improvement question asks whether the redesigned system reduces total preventable harm and cost over a relevant period, including effects outside the facility boundary.

20.6.3 Flow measurement and the paradox of local efficiency

Lean improvement benefits from explicit time and flow denominators. If a hypothetical packaging process is available for 420 productive minutes per day and genuine customer demand is 350 acceptable packets,

its takt time is 420/350 = 1.20 minutes per packet. Takt time is a pace derived from available time and demand, not a physical guarantee that every workstation can safely run at that speed. If a critical sealing and verification activity needs 1.40 minutes per packet, demanding the 1.20-minute pace without redesign creates a mismatch between work content and commercial promise. Management may need to change staffing, rearrange work, introduce a validated fixture or negotiate delivery timing, while preserving inspection and rest requirements. Faster downstream sorting does not increase the capacity of the constrained safe process (Shah & Ward, 2007).

Little's law provides a related boundary for stable flow systems: average work in process equals average throughput multiplied by average flow time, when the assumptions of consistent boundaries and long-run averages are appropriate. It does not promise that reducing inventory by decree will accelerate an unstable supply chain. Waiting may result from variation, missing release authority or materials arriving in large batches; a flow intervention should identify the mechanism rather than simply lower reported stock. A small enterprise with uncertain border lead times can require strategically placed buffers even while eliminating internal rework and unnecessary handoffs. Product release, human safety and network resilience are therefore design constraints within lean thinking, not exceptions to be quietly removed (Montgomery, 2020; Shah & Ward, 2007).

20.7 SIX SIGMA AND STRUCTURED IMPROVEMENT

20.7.1 The DMAIC logic

Six Sigma is best understood as a family of disciplined improvement practices emphasizing project selection, measurement, structured inquiry and control rather than as a universal promise of a particular defects-per-million figure. The define–measure–analyze–improve–control (DMAIC) cycle begins by establishing the problem, its customer and stakeholder consequences, process boundaries, project sponsor and measurable success criteria. Measurement then specifies the critical-to-quality characteristic, opportunity definition, baseline performance and reliability of the data collection system. Analysis evaluates variation and candidate mechanisms; improvement tests alternatives and implements a defensible change; control protects the gain with ownership, monitoring and response rules (Zu et al., 2008).

Zu et al. (2008) argue that Six Sigma introduced distinctive organizational role structures, structured improvement procedures and metric-centered discipline within broader quality-management practice. These devices can improve accountability where specialized analysis and cross-functional authority are genuinely necessary. They can also create bureaucracy, certification hierarchies and project selection bias. A 'Black Belt' label does not establish competence, just as calculating a small p-value does not by itself establish practical benefit. Managers should select improvement projects against the size and severity of harm, preventability, cost of inaction, availability of credible data and feasibility of transferring findings into durable process controls.

20.7.2 Defects, opportunities and probability assumptions

Defects per million opportunities (DPMO) is calculated as DPMO = [number of defects/(units × opportunities per unit)] × 1,000,000, provided each opportunity is explicitly defined and the counted defects correspond to that opportunity architecture. If 200 defects are recorded in 12,000 units, reporting 16,667 DPMO assumes exactly one relevant opportunity per unit and one count per defect event. This may be false: a packet may have two distinct nonconformities or multiple chances for a seal fault, and an apparent change in DPMO may arise from redefining opportunities. Defective-unit rates, defect-event frequencies and customer escape rates must therefore be presented separately.

Popular descriptions of 'Six Sigma' equate a particular sigma level with 3.4 defects per million opportunities using a conventional 1.5-sigma shift assumption. That convention is not a universal probability identity. For a centred normal distribution, six standard deviations to a specification boundary gives a different one-sided tail probability than a process assumed to drift by 1.5 standard deviations. The arithmetic is only meaningful after specifying distribution, stability, drift, opportunity count and specification. In small enterprises with sparse data, a stable and correctly measured defect fraction with confidence bounds is more defensible than a dramatic sigma score extracted from an online calculator.

20.7.3 Integrating DMAIC with experimentation and learning

DMAIC need not become a rigid sequential bureaucracy. Initial measurement may reveal that the project definition was wrong, and a proposed solution may require returning to data collection. A designed experiment that varies sealing temperature and contact time can identify technical response surfaces under controlled conditions; an observational before–after comparison cannot generally remove history or selection effects. When sample sizes are constrained, teams can start with repeatable small tests and explicit uncertainty, not fictitious statistical power. Once the improvement has been shown to operate within safe constraints, standard work, training, maintenance triggers and management review prevent the project from becoming a one-off report.

Lean and Six Sigma share concern for stable, reliable flow but differ in their common emphasis. Lean often begins with queues, work design and flow; Six Sigma often begins with variation, a critical-to-quality output and a measurable project. Neither should outrank legally required safety controls or the right to report deviations. Their joint value emerges when a team can specify what is harmful, measure it validly, change the causal mechanism and sustain the remedy without transferring losses to other functions.

20.7.4 Experimental design and the problem of generalization

When multiple process variables interact, changing one factor at a time can be misleading. Suppose seal strength depends jointly on temperature, pressure, dwell time and film type; a temperature increase that improves one film may damage another. A small factorial experiment can estimate both main effects and selected interactions if its settings are safe, randomized where practicable and replicated sufficiently to estimate variation. Factors must be chosen within engineering and regulatory limits, and measurement must be independent of the operator who favors a proposed setting. Replication across shifts, product lots and operators tests whether the result is robust rather than an artifact of one convenient trial. Montgomery (2019) emphasizes design of experiments as a method for investigating such process response structures, not a substitute for contextual expertise.

The experimental unit must be defined with care. If temperature is set once for an entire production shift, packets produced within that shift are not independent experimental assignments to different temperatures. Treating hundreds of packets as hundreds of independent replications would overstate precision. Likewise, a sequential trial may confound an apparent process improvement with operator learning or a change in material supplier. Documenting randomization, blocking, run order, calibration and outcome thresholds makes the eventual claim inspectable. A pilot that succeeds only for one product or one experienced team should not be called organizational excellence until its transfer conditions have been tested. Controls that prevent unsafe release must remain in effect throughout the experiment.

20.8 QUALITY COSTING AND BUSINESS PERFORMANCE

20.8.1 Prevention, appraisal and failure costs

Cost of quality (COQ) accounts for resources devoted to achieving adequate quality as well as losses arising from inadequate quality. A common prevention–appraisal–failure model classifies training, robust design and supplier qualification as prevention; measurement, inspection and audits as appraisal; scrap, rework and in-process failures as internal failure; and returns, warranty service, recall and post-delivery remediation as external failure (ASQ, n.d.-c). This is an analytical classification, not a declaration that every spending item is intrinsically beneficial or that all costs can be reliably recovered from ordinary ledgers. A cost may move between categories as work is redesigned. For example, an upstream automated seal test may raise appraisal costs while sharply reducing customer escapes.

An auditable COQ system should state period, product or service scope, allocation rules, data provenance and excluded costs. Internal waste requires an opportunity-cost judgment: scrap materials already consumed and extra machine hours may have different marginal economic consequences. Lost reputation, customer distress, latent injury, regulatory enforcement and future business foregone can be consequential but are rarely directly identifiable from invoice records. Reporting only easily coded rework understates failure; presenting speculative lifetime losses as certain would be equally irresponsible. Financial accounting in Chapter 22 and management costing in Chapter 23 provide complementary classification and allocation structures; quality analysis adds causal and stakeholder interpretation.

20.8.2 A conditional investment model

The constructed enterprise's monthly COQ illustration below assumes that prevention cost rises from RWF 0.90 million to 1.45 million and appraisal from RWF 0.70 million to 0.85 million following corrective action. Internal failure is assumed to decline from RWF 2.40 million to 1.20 million and external failure from RWF 1.60 million to 0.70 million. The current monthly total is RWF 5.60 million and the hypothetical post-improvement total is RWF 4.20 million, a conditional saving of RWF 1.40 million per month. Those numbers are constructed assumptions, not a measured result or a promise about Rwanda's packaging industry.

Table 20.1

Illustrative monthly cost-of-quality bridge (RWF million)

Category	Current	Conditional improved	Change
Prevention	0.90	1.45	+0.55
Appraisal	0.70	0.85	+0.15
Internal failure	2.40	1.20	−1.20
External failure	1.60	0.70	−0.90
Total COQ	5.60	4.20	−1.40

Note. Constructed illustrative data and decision rules; not observations from an identified enterprise. Source: Author's calculations and synthesis (2026).

Suppose the enterprise must spend RWF 3.00 million once on equipment adjustments, validation and training, and the monthly improvement assumptions persist for twelve months. Undiscounted first-year gross saving is RWF 16.80 million; first-year net benefit after this one-off cost is RWF 13.80 million, with a conditional simple net benefit-to-one-off-cost ratio of 4.60. This is not an internal rate of return, a risk-

adjusted net present value or a causal estimate. A conservative half-realization sensitivity produces RWF 8.40 million annual savings and RWF 5.40 million net benefit before discounting. Actual evaluation must verify that the cost categories are nonoverlapping, throughput and product mix stay comparable, new maintenance costs are included and harm reduction does not depend on suppressing complaints.

20.8.3 When quality investment is not optional

Economic appraisal cannot be the sole acceptance criterion where a product breaches mandatory safety limits or where exposure to severe harm is unacceptable. In such cases the first managerial task is containment and regulatory compliance; investment appraisal compares lawful and technically adequate alternatives. This distinction is necessary because expected monetary savings may undervalue low-frequency, high-severity events. A financial model that assigns a small cost to an unreported hazard is not neutral. Equally, limited budgets require proportionate risk prioritization and transparent staging rather than a claim that any quality expenditure must pay for itself immediately. Quality economics is a tool for responsible choice, not a license to trade away human dignity or legal protections.

20.8.4 Cash flows, discounting and the limits of value attribution

The cost-of-quality illustration in Table 20.1 describes an accounting bridge, not a cash-flow forecast. To appraise the assumed RWF 3.00 million outlay as an investment, the analyst must decide which monthly savings are incremental, cash-realizable and attributable to the intervention, whether the machine requires additional maintenance, whether working capital changes, and whether savings begin immediately. If the hypothetical RWF 1.40 million monthly net reduction were realized at each month-end for twelve months, with an illustrative 1% monthly discount rate and no terminal value, the net present value would be −3.00 + 1.40[(1 − 1.01^(−12))/0.01] = approximately RWF 12.76 million. This calculation is illustrative and represents neither an observed return nor a Rwanda market interest rate. At half the assumed monthly saving, the corresponding value is approximately RWF 4.88 million; the economic judgment is sensitive to realization, timing and scope rather than the attractiveness of the headline ratio.

The appraisal also requires a counterfactual. Some defects would have declined through ordinary maintenance or changed customer mix even without the project, while some recorded failures may rise because inspectors become more candid. A quality investment should therefore be monitored against a pre-specified comparable baseline or a suitable control process where feasible. External harms such as unsafe consumption and environmental damage must be reported separately when monetary valuation would conceal ethical constraints or rest on weak evidence. These distinctions prevent a nominal return on prevention from becoming an unsupported statement of causal effectiveness. The financial analysis should inform, but must not displace, the release authority's assessment of compliance and risk.

20.9 SERVICE QUALITY, COMPLAINTS AND CUSTOMER OUTCOMES

20.9.1 Why a service must be measured differently

Services often involve simultaneous production and use, customer participation, information exchange and heterogeneous frontline encounters. Consequently, identical technical procedures can generate different experiences and outcomes. Service quality can be considered through reliability, responsiveness, assurance, empathy and tangible signals, but these dimensions are not a universally validated metric for all cultures and sectors. Parasuraman et al. (1988) developed the SERVQUAL instrument, which compares expectations and perceptions. Cronin and Taylor (1992) challenged aspects of this approach and found support for performance-based evaluation. The dispute illustrates a deeper problem: satisfaction, attitude, objective service achievement and justice after a failure are connected but distinct constructs.

A service indicator should preserve its observational unit. 'Average response time' can be improved by ignoring difficult cases, while a high closure rate can conceal unresolved problems. For a logistics help desk, useful paired measures might include time to first meaningful response, verified case resolution, recurrence after closure, customer-reported comprehension and rights of appeal. For a bank, healthcare provider or public-facing service, security, confidentiality and lawful access may override small improvements in average handling time. Nonrespondents and customers denied service should not vanish from quality assessment merely because the existing questionnaire samples successful transactions.

20.9.2 Complaints and remedies as learning infrastructure

A legitimate complaints system makes it possible to submit a concern without retaliation, identifies the responsible recipient, acknowledges receipt, protects confidential information, distinguishes immediate remedial action from systemic correction and offers escalation where needed. ISO 10002:2018 provides complaints-handling guidance suitable for varied organizations, including small businesses; it does not determine every sector's legal obligations (ISO, 2018b). Complaints must be coded using stable categories and linked where feasible to product lot, transaction, process and resolution date. Declining complaint counts may indicate improvement, poor access, fear of reprisal or changed denominator. Triangulation with returns, repair costs, independent customer checks and frontline observations is indispensable.

Service recovery cannot universally erase failure. McCollough et al. (2000) reported that satisfaction after failure and high-quality recovery was generally lower than satisfaction from an error-free service in their experimental settings. This challenges the simplistic 'service recovery paradox' that an impressive apology makes failure desirable. Effective recovery involves acknowledgement, practical remedy, fair explanation and evidence that recurrence is being addressed. When a defect creates safety harm, customer communication must avoid language that treats compensation as proof of safety. Marketing (Chapters 13–16) owns broader market design and communication; quality management must ensure that claims are supported by actual process evidence and that complaints are converted into accountable correction.

20.9.3 Service quality in African and informal contexts

Service standards transferred from global franchises may rely on digital identity, uninterrupted internet, individual mobile-phone ownership or standardized waiting environments. Such assumptions can exclude customers whose access is mediated by shared devices, informal trade relationships, language differences or transport limitations. Quality measurement should specify whether the service is reachable, intelligible and fair to its intended user groups. In a small Rwandan enterprise, a durable notebook with unique complaint numbers, clear response deadlines and independent review of serious cases may be a better first intervention than an expensive CRM dashboard. Digitalization becomes valuable when it strengthens evidence, not merely when it modernizes appearances.

20.9.4 Service blueprints, access and outcome validity

A service blueprint links the customer's visible journey to backstage decisions and enabling systems. Consider a hypothetical Rwandan equipment-repair service: the customer reports a fault, receives an estimated appointment window, grants access, observes a repair and signs completion. Behind that journey are parts availability, technician competence, diagnosis authorization, travel scheduling, data protection, billing and complaint escalation. A short average visit may indicate efficient diagnosis or premature closure; a low number of complaints may reflect satisfaction or an inaccessible reporting route. Quality assessment therefore requires evidence at several levels: promised versus actual arrival, first-time verified resolution, recurrence within a defined window, accurate and intelligible charges, and fair treatment of customers whose locations or connectivity make service costly.

Donabedian's (1988) structure–process–outcome distinction offers a useful comparative lens for this case when carefully adapted. Having qualified technicians and test instruments is a structural condition; conducting an approved diagnostic and explaining the repair are processes; sustained operation and customer safety are outcomes. None necessarily follows from the preceding one. A well-equipped branch can perform poor diagnoses, while an excellent interaction cannot guarantee long-term equipment reliability. Managers should report relevant exclusions and nonresponse, particularly in service networks where people denied access never enter the customer survey. The implication for service quality is that a complete evidence chain is stronger than any isolated satisfaction index.

20.10 SUPPLIER QUALITY, INCOMING CONTROL AND AUDIT

20.10.1 Quality extends across organizational boundaries

An enterprise cannot fully control its finished product if purchased materials, outsourced processes and transport conditions are treated as externalities. Supplier quality begins before a purchase order: a specification must identify critical functional characteristics, acceptance methods, evidence requirements, traceability, packaging, storage and responsibilities if material fails. Procurement owns fair supplier selection and contractual governance (Chapter 18); quality management owns the technical assurance that purchased inputs and outsourced work conform to legitimate product and process requirements. The supplier with the lowest unit price may create high inspection, rework, recall or continuity costs, making total landed and lifecycle consequences relevant to selection.

Incoming inspection is risk-weighted. Low-risk standardized stationery may be accepted under supplier history and sample checks; a high-risk food ingredient or safety-critical component may require identity verification, documentation, accredited test evidence, controlled storage and possibly intensified testing. Sampling can reduce inspection expenditure, but acceptance of a lot under a statistical sampling plan never logically proves that every unit conforms. The producer must understand the producer's and consumer's risks embedded in the plan and apply regulatory controls independently of commercial convenience. Measurement traceability matters because a certificate produced by an unverified laboratory may offer less assurance than its professional appearance suggests.

20.10.2 Audits, corrective action and traceability

Supplier audits should be designed around the risks that contracts and test results cannot reveal: process changes, record integrity, staff competence, subcontracting and segregation of defective material. A supplier scorecard might combine defect frequency, severity, on-time corrected delivery, repeat nonconformity, traceability exercise performance and documented corrective-action timeliness. Each indicator needs a denominator and an independent means of challenge. An average supplier defect rate can conceal a single unacceptable safety incident, while weighting all suppliers equally can obscure exposure. Risk priority should consider consequence, probability, detectability, substitute availability and capacity to recover, not a decorative total score alone.

Traceability means being able to follow an identified input or output through relevant production and distribution stages. A practical test asks the enterprise to identify all affected finished-product lots and customers from one suspect input batch, and then to trace a finished lot backward to relevant ingredients, dates, suppliers and equipment records. The critical measure is not the existence of software but completeness and time to credible identification. Where traceability is deficient, management must broaden containment and disclose uncertainty. Customers and regulators deserve accurate information even when disclosure creates near-term commercial losses.

20.10.3 Supplier development and regional standards

A defensible supplier improvement programme provides explicit standards, timely technical feedback, root-cause collaboration, access to attainable testing and proportionate corrective deadlines. Abruptly excluding micro-suppliers that lack expensive certificates may protect a large buyer's formal audit score while reducing local participation and competition. Conversely, affordability cannot justify accepting unsafe inputs. Rwanda's quality infrastructure includes standardization, metrology, testing, certification and related accreditation functions, and RSB provides product and systems certification information (RSB, 2018, n.d.-b). Under East African regional trade, mutual recognition and harmonized standards can reduce repeated testing only where the relevant instruments, recognized scope and conformity requirements actually apply. Supplier development is thus a shared task for procurement, operations, legal compliance and quality specialists rather than a unilateral demand for paperwork.

20.10.4 Laboratory competence and cross-border evidence

Credible quality decisions require clarity about the organizations that generate test results. ISO/IEC 17025:2017 establishes requirements for competent, impartial and consistently operated testing and calibration laboratories, while ISO 9001:2026 concerns the management system of an organization more generally. An ISO 9001 certificate held by a supplier is not interchangeable with laboratory accreditation for a particular analytical method; conversely, a laboratory's accredited scope may not cover every characteristic displayed on its price list (International Organization for Standardization/International Electrotechnical Commission [ISO/IEC], 2017; ISO, 2026b). A Rwandan buyer should therefore verify the method, uncertainty where material, tested matrix, sampling custody, calibration traceability, accreditation scope and whether a report refers to the delivered lot. When a failed result is disputed, independent retesting should be specified with an evidential chain before contractual sanctions are imposed.

Cross-border trade adds a further institutional problem: technical equivalence does not automatically imply legal recognition of another country's certificate. Recognition may depend on harmonized standards, accreditation arrangements, destination requirements, shipment identity and competent-authority decisions. For smaller enterprises, repeated testing can be costly and can discourage regional participation. However, waiving needed safety evidence to accelerate a shipment shifts risk onto consumers. The practical solution is to document applicable requirements before contracting, select recognized technical services and retain lot-level evidence that is usable across procurement, border clearance, customer assurance and any later corrective action. Quality management thus converts regional integration from an abstract promise into a disciplined system of verifiable claims.

20.11 QUALITY CULTURE, LEADERSHIP AND WORKER AGENCY

20.11.1 The behavioural foundations of error disclosure

No procedure can compensate for organizational incentives that punish the people who report errors. A quality culture is demonstrated when leaders act upon unwelcome evidence, protect credible concern-raising, allocate resources for prevention and accept personal responsibility for management-created conditions. Deming's critique of fear is especially relevant where employees are graded by numbers they cannot control (Deming, 1986). A worker who sees a seal failure must know how to stop or flag output, and should not face retaliation for using that authorization. Supervisors must distinguish honest human error, reckless disregard of established protections and deliberate falsification without treating every deviation as grounds for blame.

Training creates capability only when instruction is reinforced through usable tools, time and authority. Requiring a worker to execute a ten-step verification procedure within a five-step time allowance makes later error predictable. Likewise, rewarding managers solely for dispatch quantity can encourage them to bypass protective holds. The HR systems developed in Chapters 9–12 and the governance mechanisms of Chapter 8 are therefore integral to quality assurance. Worker participation should include practical design review, accessible safety information, opportunities to test proposed changes and feedback on management's response, rather than a suggestion box whose entries are never considered.

20.11.2 Leadership, audit candour and difficult trade-offs

High-quality reporting requires credible independence. A quality manager asked to sign off a batch with missing safety records should have a documented escalation route that cannot be overridden by a salesperson's commercial target. However, independence should not become an unaccountable veto: the evidence, requirement and basis for disposition must remain explicit and contestable by competent review. Quality leaders must also guard against managerial appropriation of successes and individualization of failures. Many defects arise from accumulated decisions on procurement specifications, understaffing, equipment maintenance and unrealistic deadlines rather than one visible operator's conduct.

A defensible quality culture should be assessed through triangulation: willingness to raise concerns, near-miss reporting, time to resolve serious findings, repeat defects, documented leadership responses and whether the same standards apply to favored and unfavored staff. Survey scores alone are insufficient, particularly if anonymity is uncertain. A rise in reports may indicate improved speaking-up conditions rather than worsening actual quality; a fall may signal intimidation. Interpretation requires local institutional knowledge and temporal comparison, not automatic ranking of teams by reported incidents.

20.11.3 Justice, power and the credibility of quality evidence

A quality system is also a distribution of epistemic authority: certain people may define the defect, approve the sampling method, interpret anomalies or authorize release. That distribution creates a risk of systematic silence when the employees who observe failures lack practical power to challenge management. A quality culture cannot be established from staff satisfaction scores alone; it requires evidence that credible reports are investigated, serious hazards are escalated irrespective of commercial consequences and those adversely affected can seek remedy. Transparency must be reciprocal. Workers should receive enough information about why an incident matters, customers should not be misled about testing or certification scope, and managers should expose their own process-design assumptions to review. Institutional protections for reporting should be evaluated against relevant local law and actual workplace practice, not inferred from the presence of a policy (Deming, 1986; Hackman & Wageman, 1995).

Fairness does not require eliminating individual accountability. Deliberate falsification, reckless circumvention and innocent human error have different moral and organizational implications, yet these distinctions cannot be determined from a failed output alone. Investigators should examine workload, conflicting instructions, equipment condition, incentives and opportunities for competent action before attributing culpability. Severe customer harm may require immediate containment independently of whether misconduct is proven. This separation of protective action, technical explanation and employment judgment prevents a search for blame from contaminating causal inquiry. It also connects quality management to employee voice, governance, labour relations and customer protection without transferring those subjects out of their designated chapters.

20.12 DIGITAL QUALITY, ARTIFICIAL INTELLIGENCE AND SUSTAINABLE SYSTEMS

20.12.1 Sensors, traceability and predictive quality

Digital quality systems can join machine signals, test results, inventory identifiers and customer complaints into a more rapid evidence chain. Sensors may monitor temperature, torque, weight, humidity or seal integrity continuously rather than relying on end-of-line sampling. Computer vision may help classify visible defects, and predictive models may identify combinations of conditions associated with failure. Yet useful deployment requires stable identifiers, validated measurement methods, maintenance, data provenance, calibrated thresholds and procedures for checking both false positives and false negatives. A highly accurate-looking dashboard can become systematically misleading when sensors drift, product mix changes or an algorithm is trained predominantly on easy defects (Mahin et al., 2024).

For an AI-assisted visual-inspection model, validation should use held-out observations representative of equipment, lighting, product variants and realistic defect severities. A confusion matrix must distinguish sensitivity to dangerous defects from overall accuracy. In a production stream with rare critical faults, a model can achieve very high headline accuracy by mostly predicting 'good'. Performance should be reported with prevalence, precision, recall for each safety-critical class, false-release rate, sample definition and uncertainty. Drift monitoring, documented human override, version control and safe fallback procedures are necessary. NIST's voluntary AI Risk Management Framework offers a useful governance vocabulary, but its existence does not certify an enterprise's model or displace domestic law (NIST, 2023).

20.12.2 Cybersecurity, data governance and automated decision rights

Quality data now depend on software configurations, access controls, networks and third-party providers. Unauthorized changes to acceptance limits, fabricated sensor readings, manipulated lot histories or unavailable systems can make a perfectly functioning production line unsafe. Quality management should therefore coordinate with cybersecurity and information systems (Chapters 28–29) on role-based permissions, change logs, backup and recovery, integrity controls, supplier-software assurance and a tested manual fallback for safety-critical release decisions. 'Human in the loop' is meaningful only when the reviewer has sufficient time, competence, contextual information and genuine authority to reject automated recommendations.

Digital capability also has opportunity costs. Buying cameras or a cloud platform before fixing inaccurate defect definitions and broken process ownership may scale unreliable decisions. A staged intervention can begin with a standardized check sheet, calibration records and lot identifiers, then introduce low-cost electronic capture, then test a sensor or model against a defined baseline. The governing criterion is incremental evidence of reliable decisions at acceptable cost and risk, not the sophistication of the purchased interface. A well-governed manual process may outperform poorly controlled automation in a low-volume enterprise.

20.12.3 Quality, resource efficiency and environmental limits

Quality failure creates environmental burdens through scrap, additional transport, energy-consuming rework, premature replacement and discarded products. Prevention can therefore align quality with resource productivity. Yet minimizing visible waste inside a factory is not sufficient: an easy-to-manufacture package may generate difficult-to-recover post-consumer waste. A relevant life-cycle perspective examines materials, energy, safe use, durability, repairability and end-of-life effects, while keeping quality compliance and environmental claims analytically separate. ISO 14001:2026 provides a current environmental-management framework distinct from ISO 9001:2026, and its use does not establish a particular product's environmental superiority (ISO, 2026d).

Rwanda's IBES 2024 identifies uneven access to capabilities and reports that 62.6% of manufacturing enterprises surveyed did not fully use installed capacity, while its measures of waste management and solar-energy use have their own survey definitions (NISR, 2026). These data do not measure quality defects, and no causal connection to quality performance should be inferred. They do suggest practical constraints on investment, process stability and capability development that an imported 'Quality 4.0' blueprint may underestimate. Responsible transformation requires anticipating energy reliability, repair services, technician training, data ownership and the environmental cost of discarded electronics as well as potential efficiency gains.

20.12.4 Rare-defect detection and the economics of false assurance

A simple numerical example shows why automated visual inspection needs class-specific evaluation. Suppose 10,000 hypothetical units contain 100 truly nonconforming items and 9,900 conforming items. An inspection model identifies 90 of the 100 nonconforming units but also rejects 495 conforming units. Sensitivity is 90/100 = 90%, specificity is 9,405/9,900 = 95%, and precision among the 585 flagged items is 90/585 = 15.4%. Overall accuracy is (90 + 9,405)/10,000 = 94.95%, yet ten defective units are falsely released, including potentially critical defects. These counts are constructed, not reported results for an AI product. The example demonstrates why a seemingly strong accuracy number does not establish safety, especially when prevalence is low. A threshold that raises sensitivity may increase costly false rejection; the appropriate threshold depends on consequences, independent verification and release rules.

Before procurement of such a system, managers should specify the defect taxonomy, acceptable false-release exposure, data provenance, representative lighting and product conditions, human review capacity, fallback method and surveillance for model drift. Evaluators should use separate development and test sets, report performance by class and production condition, and avoid leaking nearly identical images of one defective lot across training and evaluation. System-wide costs include camera maintenance, calibration, cybersecurity, rejected good product, manual adjudication, retraining and vendor dependence. Digital inspection is therefore a potential source of quality capability, not independent proof of it (Mahin et al., 2024; NIST, 2023).

20.13 RWANDA APPLICATION: AN AUDITABLE FOOD-PACKAGING IMPROVEMENT CASE

20.13.1 Institutional context and evidentiary boundaries

Rwanda's National Institute of Statistics reports approximately 278,060 non-agricultural business enterprises in its 2024 Integrated Business Enterprise Survey, with an informality share of at least 87% under the survey's definitions (NISR, 2026). This makes a quality programme that presupposes a full-time statistical engineer, enterprise resource-planning suite and laboratory on every site implausible as a default for the economy. RSB's activities in standards, conformity assessment, metrology and certification establish a national institutional setting for credible requirements and testing; its product certification programme identifies the Standardization Mark (S-Mark), while the Made in Rwanda mark is governed by its own criteria (RSB, n.d.-b, n.d.-c). Such marks have specified scopes and should not be presented as unrestricted endorsements of every business claim. The institutional facts provide context, not observed quality outcomes for the hypothetical enterprise below.

The teaching case concerns 'Kigali Valley Foods', a fictitious enterprise that manufactures shelf-stable packaged snacks for retail and institutional buyers. Assume production of 12,000 packets in a month, under one process scope, with a common product specification, six recording categories and consistent procedures for capturing defect events. The case provides two different constructed data systems: a production-wide check-sheet of 200 recorded defect events, where one packet may contribute multiple

events, and twelve sequential samples of 200 packets each, for which the inspector records whether each packet has at least one defined nonconformity. These cannot be equated without item-level linkage. Neither system has been observed in Rwanda; their value is methodological transparency and an opportunity to examine what the measures can and cannot establish.

20.13.2 Reproducible Pareto analysis

Table 20.2

Illustrative defect-event Pareto analysis for 12,000 packets

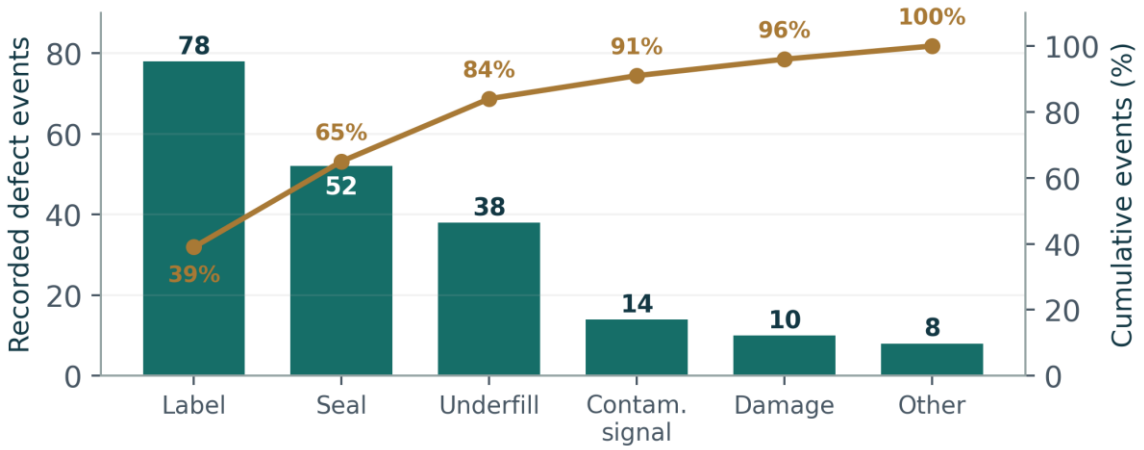
Defect-event category	Events	Share of 200	Cumulative share
Incorrect or unreadable label	78	39%	39%
Weak or incomplete seal	52	26%	65%
Underfilled packet	38	19%	84%
Potential contamination indicator	14	7%	91%
External package damage	10	5%	96%
Other documented events	8	4%	100%
TOTAL	200	100%	100%

Note. Constructed illustrative data and decision rules; not observations from an identified enterprise. Source: Author's calculations and synthesis (2026).

The check-sheet yields 78/200 = 39% for labeling and 52/200 = 26% for seals; together those event types account for 130/200 = 65%. Adding underfill brings the cumulative event share to 84%. These percentages permit a ranking of observed defect-event frequency, not a conclusion that 84% of defective packets arise from three causes, nor that every counted event has equal severity or cost. The category labeled 'potential contamination indicator' is lower frequency yet may warrant immediate containment and formal safety review, depending on the hazard and evidence. A sound priority rule therefore considers frequency, severity, customer exposure, detectability and remedy cost separately. The familiar 80/20 phrase is a heuristic, not a theorem governing this dataset or a reason to ignore serious infrequent events.

Figure 20.3

Pareto chart of constructed defect-event frequencies and cumulative percentages



Note. Original explanatory figure synthesizing the cited concepts; all case values, where shown, are hypothetical.

20.13.3 Process-stability evidence

Table 20.3

Sequential sample data for p-chart analysis, all groups n = 200

Sample	Nonconforming units	Observed fraction
1	7	0.035
2	9	0.045
3	8	0.040
4	8	0.040
5	10	0.050
6	6	0.030
7	9	0.045
8	11	0.055
9	7	0.035
10	8	0.040
11	8	0.040
12	23	0.115

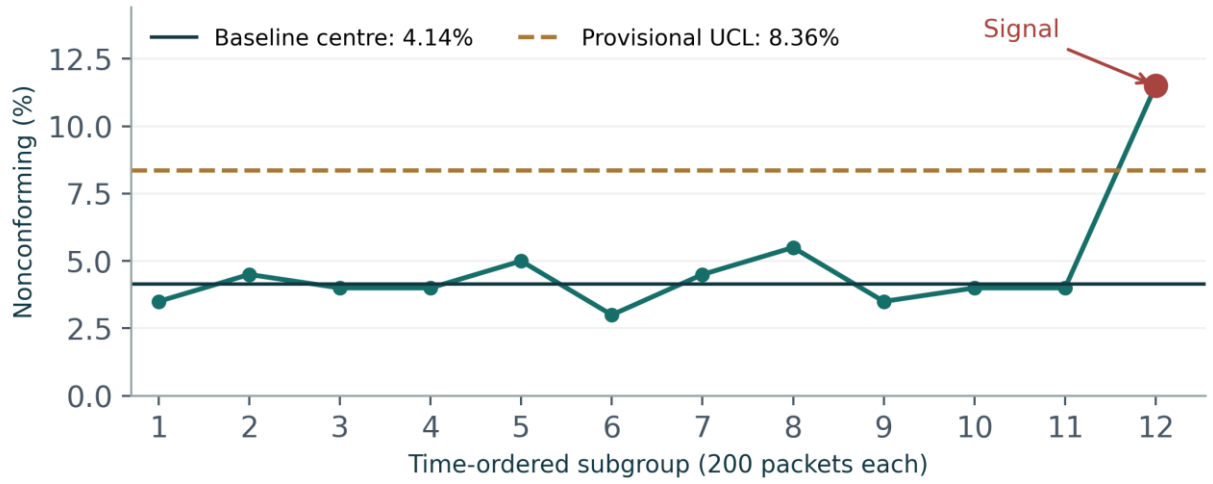
Note. Constructed illustrative data and decision rules; not observations from an identified enterprise. Source: Author's calculations and synthesis (2026).

The eleven preliminary subgroups contain 91 nonconforming packets among 2,200 inspected. Their pooled proportion is 4.136%, with the binomial baseline upper three-sigma limit about 8.361% and a lower limit truncated at zero. The twelfth subgroup is 11.5%, beyond that provisional upper limit. It is therefore

reasonable to initiate a special-cause investigation and quarantine potentially affected work if the defect class warrants it. The result is not a formal declaration that all prior production was stable: eleven baseline points provide limited evidence, and dependence within a shift, sampling bias or inconsistent classification may invalidate binomial assumptions. A formal operational system would ordinarily need a longer representative baseline and clearly specified supplementary run rules.

Figure 20.4

Illustrative p-chart with eleven baseline samples and one investigative signal



Note. Original explanatory figure synthesizing the cited concepts; all case values, where shown, are hypothetical.

20.13.4 Causal inquiry and control of unsafe release

A fishbone team can first classify hypotheses under materials, equipment, method, measurement, people and environment. For weak seals, candidate mechanisms might include variable film thickness, worn heating components, inadequate dwell time, unverified gauge calibration, inconsistent handoff instructions and heat or humidity effects. The case does not supply experimental evidence that any is the actual cause. The team should record the film lot, sealer, temperature setting, time of each event and inspection operator; assess agreement on seal classification; and test carefully controlled settings against qualified acceptance criteria. Where product safety is implicated, pending stock should remain under an authorized disposition rule rather than be released because the average defect score appears acceptable.

The corrective-action plan would initially contain affected product, perform independent seal verification and protect customers where an unsafe escape is plausible. It would then investigate and test a causal hypothesis, formally approve the process change, retrain relevant roles, verify effectiveness with independent sampling and monitor recurrence. Releasing staff from impossible cycle-time targets may be part of the change if workload is causal. The case does not demonstrate the eventual improvement: the post-change values in Table 20.1 remain conditional forecasts requiring measured validation. This distinction keeps the pedagogical exercise from becoming fabricated evidence of a successful intervention.

Table 20.4

Illustrative corrective-action and evidence register

Issue or decision	Accountable role	Required evidence	Closure test
Suspect lots and containment	Quality release owner	Lot list, safety assessment, disposition authorization	No unapproved lot released
Seal root cause	Operations and quality analysts	Measurement agreement, lot/shift trace, controlled test	Cause hypothesis survives verification
Approved process change	Operations manager	Change record, revised specification, competence evidence	Authorized trial meets specification
Sustained performance	Independent quality reviewer	Time-series samples, escape records, complaint trends	Stable performance and no new safety signal
Supplier countermeasure	Procurement and supplier quality	Material tests and supplier corrective action	Verified specification and repeat-lot conformity

Note. Constructed illustrative data and decision rules; not observations from an identified enterprise. Source: Author's calculations and synthesis (2026).

20.13.5 Integrated decision memo and staged adoption

A responsible manager should distinguish immediate customer protection from the financial question of financing durable process change. If the RWF 3.00 million initial investment is necessary for a safety-critical remedy, the organization must pursue a lawful, technically adequate alternative regardless of whether the first-year economic model is attractive. If it concerns discretionary quality enhancement beyond compliance, the model in Table 20.1 supports a staged decision: approve an instrument and process-validation pilot, define acceptance criteria, verify supplier inputs, train staff, then release recurring expenditure when reliable evidence shows improvement. Managers should separately review cash impact, working-capital needs and the risk that production growth changes defect exposure.

A suitable improvement scorecard records a small balanced set of measures rather than a single 'quality score': sampled nonconforming proportion with control-chart status, safety-critical escapes, cost of poor quality with scope, complaint resolution and recurrence, supplier-lot conformity, audit actions overdue, and validated waste or energy per usable packet. Every measure should include a clear denominator, accountable owner, period, source and corrective trigger. Public communication must not claim that a simulation establishes actual enterprise performance; transparent reporting is itself a quality obligation.

Table 20.5

Proposed quality-improvement scoreboard and decision rules

Indicator	Numerator and denominator	Interpretive safeguard	Action trigger
Unit nonconformance	Defective audited packets / audited packets	Risk-weighted sampling	Special-cause signal or unsafe unit
Escape rate	Confirmed customer escapes / delivered packets	Complaint access affects count	Any critical escape and adverse trend
Cost of poor quality	Internal + external failure RWF / monthly sales RWF	Exclude speculative values from ledger measure	Adverse trend with stable scope
Supplier conformity	Rejected critical lots / tested critical lots	Supplier and input stratification	Critical failure or worsening trend
Corrective action effectiveness	Verified effective actions / due actions	Evidence, not closure signatures	Repeated issue or overdue critical action
Waste intensity	Discarded material kg / accepted packets	Include off-site transferred waste	Increase without product-mix explanation

Note. Constructed illustrative data and decision rules; not observations from an identified enterprise. Source: Author's calculations and synthesis (2026).

20.14 INTERNATIONAL COMPARISON: WHAT CAN AND CANNOT TRAVEL

20.14.1 Toyota as a documented institutional comparator

Toyota describes its production system through just-in-time and jidoka, with foundational routines developed through repeated experimentation and extended across facilities and suppliers over decades (Toyota Motor Corporation, n.d.). The relevant lesson for a Rwandan enterprise is not that an inexpensive visual board recreates Toyota's results. It is that quality, process flow, employee judgment, equipment design and supplier development have to be jointly configured. Jidoka is especially instructive as a principle of detecting abnormal conditions and interrupting unsuitable production rather than multiplying inspection after the fact. Toyota's account is an organizational self-description, not independent proof of every claimed economic or labor outcome, and it cannot establish the feasibility of identical practices in Kigali.

There are identifiable boundary conditions. Toyota's scale, supplier density, industrial history, engineering capabilities, investment horizon and employment institutions differ from those of microenterprises working with imported inputs and thin margins. A low-inventory strategy that is coherent within a resilient supply network may be hazardous where border delays and machine spares are uncertain, as Chapter 19 demonstrated. Conversely, a small enterprise may be able to change an operating instruction or experiment with a fixture more quickly than a complex multinational. Comparative transfer should focus on the underlying causal mechanism, identify the necessary complementary capabilities and test local feasibility rather than copy tools by name.

20.14.2 Certification research and the problem of selection

The U.S. certification study by Corbett et al. (2005) provides evidence that certified manufacturers subsequently exhibited favorable financial performance relative to selected comparators, while showing sensitivity to matching choices. This supports taking management systems seriously without claiming that the certificate itself is the unique causal agent. Firms that pursue certification may already have better

management, more demanding customers or access to larger capital markets. The study's industrial, national and historical boundaries also restrict direct inference to Rwanda's informal and service-intensive enterprise population. Research of this kind is most valuable when it encourages careful counterfactual evaluation instead of mechanical promises about returns on certification fees.

20.14.3 Rwanda's standards ecosystem and regional markets

RSB's quality policy defines the country's national quality infrastructure through standards, measurement, conformity assessment and accreditation-related arrangements, while its public guidance explains product certification and technical services (RSB, 2018, n.d.-b). For exporters and suppliers serving institutional clients, credible testing and recognized standards can be commercially important, especially where buyers have limited direct knowledge of product quality. Yet a certificate can exclude capable small suppliers where costs are disproportionate; recognition across borders depends on the actual scope and institutional arrangements. Improvement programmes should therefore combine clear risk requirements with accessible laboratory and metrology services, supplier learning and fair recognition of reliable evidence. Formalization is a means of building warranted confidence, not an adequate substitute for it.

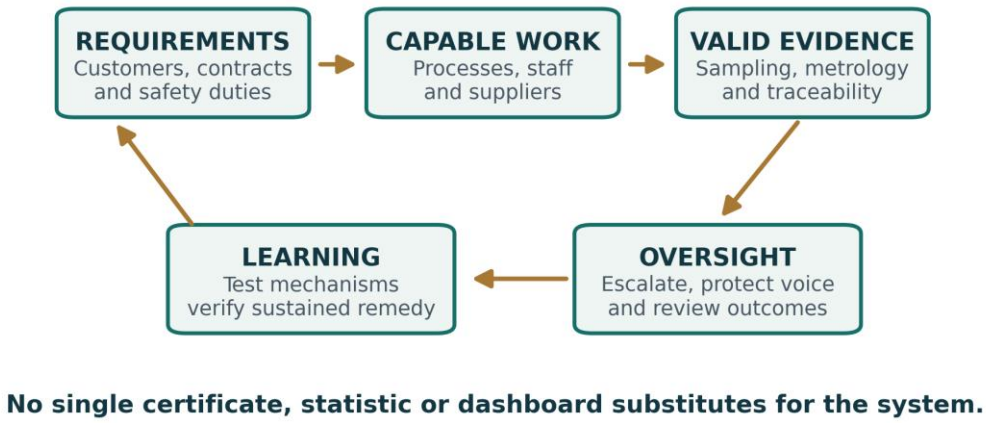
20.15 INTEGRATIVE DISCUSSION, CRITICAL SYNTHESIS AND RESEARCH DIRECTIONS

20.15.1 A conditional theory of accountable quality capability

Quality emerges through a sequence of connected but distinct propositions. First, legitimate requirements translate heterogeneous needs, contractual commitments and non-negotiable safety duties into testable characteristics. Second, production and service systems can only achieve those characteristics reliably if process variation is understood and material, human and informational resources are capable. Third, monitoring is useful only where measurement is valid and personnel can act on the evidence without fear. Fourth, corrective action requires a causal account whose implications have been tested, not a narrative that ends at 'operator error'. Fifth, the gains must survive supplier variation, market change, managerial incentives and digital-system failure. These links form an analytical model, not empirical proof of a new causal system.

Figure 20.5

An integrated enterprise architecture for accountable quality capability



Note. Original explanatory figure synthesizing the cited concepts; all case values, where shown, are hypothetical.

The model has important rival explanations. A firm's quality gains may reflect more favorable customers, newer equipment or a demand shock rather than its formal QMS. Customer satisfaction may rise through reduced expectations rather than better service. A fall in recorded defects may follow more permissive inspection, not safer production. A successful pilot may fail when expanded to other shifts or suppliers. These alternatives should be specified before data collection. Quality research is especially vulnerable to same-source questionnaire bias, survivor selection, inadequate counterfactuals and changes in the definitions of 'defect' or 'success' across time (Flynn et al., 1994; Kaynak, 2003).

20.15.2 Cross-functional consequences and responsibilities

Quality is not the property of an inspection office. Marketing and product managers determine which promises customers can reasonably expect, while operations makes the promised specification technically feasible. Procurement influences input variation and traceability, logistics protects condition in storage and transit, and finance determines the feasibility of prevention investment and the reliability of cost classifications. HR and organizational leadership structure learning, reporting protection and workload, while information systems secure the records underlying release and recall decisions. Governance is required where local commercial pressures threaten safety or evidence integrity. This interdependence is the same enterprise-system logic established in Chapter 1, expressed here as a set of quality assurance decision rights.

The relation between quality and sustainability deserves particular caution. Fewer defects may reduce material use, but a production change could increase electricity consumption or generate nonrecyclable packaging. A cost-saving automation may reduce human judgment where it is most essential; a customer-driven customization programme may compromise safe standardization. These are not arguments against improvement. They require evaluating enterprise performance with financial, technical, environmental and ethical boundaries made explicit, instead of treating one trend as conclusive.

20.15.3 Testable propositions and a realistic research agenda

One research question concerns whether locally adapted quality practices produce greater improvements in defect escape rates than formal certification alone among Rwandan small enterprises after baseline process capability, sector, customer composition and managerial resources are accounted for. A second concerns whether protected employee reporting mediates the relation between formal procedures and detection of safety-critical anomalies. A third asks whether supplier-development support reduces quality disparities between informal and formal suppliers without increasing unsafe acceptance. A fourth examines the incremental predictive value and fairness of AI-assisted inspection against simpler instrumented process controls after allowing for machine drift and rare defects.

A defensible programme would first develop validated multilingual measures and stable defect taxonomies, identify relevant populations and sampling frames, and predefine product-level or service-level outcomes. Where feasible, staggered implementation with comparison enterprises could support difference-in-differences analysis, provided parallel-trend assumptions are examined; a randomized phased rollout may be possible for low-risk non-mandatory interventions. Mixed documentary, operational and employee-voice evidence could explain mechanisms overlooked by dashboard statistics. For low-frequency severe hazards, prospective monitoring and qualitative process tracing may be more informative than underpowered inferential tests. Researchers should publish deviations, measurement changes, negative findings and scope restrictions rather than treating every improvement project as proof of the favored theory.

20.15.4 Boundary conditions and a programme for independent evaluation

The integrated framework is not a claim that all enterprises should maximize formalization, inspection intensity or automation. It has at least four boundary conditions. First, high-severity hazards create binding release constraints that cannot be offset by favorable averages. Second, institutional access to competent testing and legal remedy varies, shaping what evidence can be produced at reasonable cost. Third, operations with infrequent outputs, nonstationary demand or bespoke services may not support conventional control-chart assumptions without substantial adaptation. Fourth, reported improvements can be artifacts of changed classification, sampling or complaint access. For these reasons, the appropriate research object is a specified quality mechanism operating within a defined institutional and technological configuration, not 'quality maturity' as an abstract quantity.

An independent evaluation should prospectively register outcomes and harms, retain original defect definitions, measure reporting access, and establish the timing and content of changes in process controls. Where comparable sites exist, a staggered implementation may offer stronger evidence than an uncontrolled before–after graph, although spillovers and selection may remain. A credible mixed-method programme would combine audit traces, direct process observation, financial ledgers, employee accounts and independently collected customer outcomes. Its normative safeguards should preclude randomizing mandatory safety protections away from any unit. Publication should include null or adverse findings, limitations of the statistical model and a clear description of which conclusions are theoretical, descriptive or causal. The same discipline is needed to evaluate the framework advanced in this chapter.

20.16 CHAPTER SYNTHESIS, TERMINOLOGY AND LEARNING RESOURCES

20.16.1 Synthesis and conclusion

Quality is best understood as a justified relationship between requirements, actual characteristics, observed consequences and accountable processes. Historical traditions from Shewhart, Deming and Juran

established that prevention and system learning are distinct from sorting defective output. Current management-system standards help organize responsibilities, but no certificate can substitute for technically competent verification. Control charts can identify signals worth investigation, whereas specification limits determine whether outputs meet requirements. Pareto analysis can rank events, but cannot turn correlation into a causal explanation or frequency into a measure of moral severity. Lean and Six Sigma provide complementary means of changing flow and variation, conditional on valid data and worker participation. Cost-of-quality analysis demonstrates why prevention can be economically attractive, while safety and legal duties impose limits that must not be reduced to expected profit. Services, suppliers, digital systems and environmental impacts extend quality across organizational boundaries. Taken together, the chapter's central finding is conceptual: operational excellence is a maintained capacity for honest, evidence-based and correctable performance, not a permanent status established by a dashboard or award.

20.16.2 Key terminology

Table 20.6

Selected chapter glossary

Term	Operational meaning
Quality	Degree to which relevant characteristics fulfil justified requirements within a specified context
Conformance	Agreement of an observed characteristic with an explicit requirement
Quality assurance	Planned confidence-building controls concerning the capacity to meet requirements
Quality control	Activities that verify or control the conformity of results
Common cause	Routine process variation within a provisionally stable generating system
Special cause	A distinguishable disturbance or pattern not adequately explained by the baseline model
Control limit	Statistically estimated process-behaviour boundary, distinct from customer specification
Nonconforming unit	A unit with one or more defined nonconformities
Defect event	A counted occurrence of a defined nonconformity, with multiple possible in one unit
Corrective action	Action intended to address causes and reduce recurrence of nonconformity
DMAIC	Define, measure, analyze, improve and control sequence for structured projects
Cost of quality	Prevention plus appraisal plus internal and external failure expenditures
Traceability	Ability to identify and follow relevant history, inputs or distribution pathways
Operational excellence	Evidence-supported, sustained and responsible performance across relevant dimensions

Note. Definitions synthesized from cited quality-management literature and standards; wording adapted for teaching.

20.16.3 Questions and applied tasks

The foundational discussion exercise requires a comparison of conformance, fitness for use, statistical control, safety and customer satisfaction, using one food product and one digital service. Readers should

explain how a process can be statistically stable while routinely generating unacceptable outputs, and identify the distinct evidence required for each of these quality claims.

In the applied statistical exercise, readers reproduce every percentage and cumulative share in Table 20.2 and develop a properly labelled two-axis Pareto chart. They then establish a corrective-action priority that distinguishes the frequency and seriousness of defects. From Table 20.3 they must calculate a provisional p-chart centre line and three-sigma limits using only samples 1–11, interpret sample 12 and explain why the adverse observation should not be silently absorbed into a newly calculated baseline.

The integrative assessment is a 2,000-word memorandum to the hypothetical board of Kigali Valley Foods. The memorandum should defend immediate containment, cause investigation, measurement verification, supplier engagement, process trials, protection of employee reporting, customer communication, resource decisions, financial assumptions and follow-up verification. It should compare a proportionate paper-based management system with an instrumented digital option using explicit criteria for safety, operating feasibility, cash flow, accessibility and environmental responsibility. The case data must be labelled as constructed rather than observed enterprise results.

The advanced research task asks students to prepare a preregisterable observational design examining quality-system changes among Rwandan small manufacturers. A defensible protocol will specify inclusion criteria, baseline and follow-up outcomes, a plausible comparison strategy, threats to identification, power and access constraints, protections for workers and customers, and a reproducibility plan. The research justification should explain why a cross-sectional survey of perceived total-quality maturity cannot by itself establish an effect on physical product defects.

20.16.4 Educator guidance and assessment rubric

An adequate foundational answer separates output conformance from user experience and safety, and recognizes that control limits describe process behaviour rather than acceptance criteria. For the applied task, check the exact cumulative sequence 39%, 65%, 84%, 91%, 96% and 100%; the preliminary baseline is 91/2,200 = 0.04136, UCL approximately 0.08361 and LCL zero, so sample 12 at 0.115 generates a provisional special-cause signal. Good work identifies the small number of baseline samples, the possibility of non-independent inspected units and the need to verify measurement. For the financial task, baseline and conditional post-intervention COQ are RWF 5.60 million and 4.20 million per month respectively, with RWF 1.40 million difference, RWF 16.80 million annualized gross conditional savings and RWF 13.80 million first-year net after RWF 3.00 million initial costs, prior to discounting. A sophisticated answer does not confuse that forecast with experimentally established benefit.

A suitable advanced rubric allocates 25% to theoretical and technical precision, 25% to evidence and calculation validity, 20% to practical and cross-functional feasibility, 20% to ethical and safety governance, and 10% to scholarly clarity and source integrity. Work claiming a certified process cannot fail, ignoring a critical hazard because of low count, manipulating denominators or presenting illustrative data as real should not achieve the highest band irrespective of rhetorical quality.

20.16.5 Further reading and publication note

For foundational accounts of quality and systemic learning, Deming (1986), Juran (1988), Shewhart (1931) and Hackman and Wageman (1995) warrant close comparative reading. The empirical literature should be approached through Flynn et al. (1994), Powell (1995), Kaynak (2003) and Zu et al. (2008), with attention to construct validity and causal identification. Montgomery (2019, 2020) provides advanced experimental and control-method treatment; the NIST/SEMATECH handbook offers accessible technical checks. The current ISO standards, the official certification-transition guidance, and dated NISR and RSB publications

should be consulted for operational and institutional application. DOI registration and external peer review remain publisher-controlled processes and should be recorded in the publisher's release record rather than assumed from the presence of a proposed identifier in the manuscript.

20.17 REFERENCES

American Society for Quality. (n.d.-a). *History of quality*. https://asq.org/quality-resources/history-of-quality

American Society for Quality. (n.d.-b). *Seven basic quality tools*. https://asq.org/quality-resources/seven-basic-quality-tools

American Society for Quality. (n.d.-c). *Cost of quality*. https://asq.org/quality-resources/cost-of-quality

Corbett, C. J., Montes-Sancho, M. J., & Kirsch, D. A. (2005). The financial impact of ISO 9000 certification in the United States: An empirical analysis. *Management Science*, *51*(7), 1046–1059. https://doi.org/10.1287/mnsc.1040.0358

Cronin, J. J., Jr., & Taylor, S. A. (1992). Measuring service quality: A reexamination and extension. *Journal of Marketing*, *56*(3), 55–68. https://doi.org/10.1177/002224299205600304

Crosby, P. B. (1979). *Quality is free: The art of making quality certain*. McGraw-Hill.

Deming, W. E. (1986). *Out of the crisis*. Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Donabedian, A. (1988). The quality of care: How can it be assessed? *JAMA*, *260*(12), 1743–1748. https://doi.org/10.1001/jama.1988.03410120089033

Flynn, B. B., Schroeder, R. G., & Sakakibara, S. (1994). A framework for quality management research and an associated measurement instrument. *Journal of Operations Management*, *11*(4), 339–366. https://doi.org/10.1016/S0272-6963(97)90004-8

Garvin, D. A. (1987). Competing on the eight dimensions of quality. *Harvard Business Review*, *65*(6), 101–109. https://hbr.org/1987/11/competing-on-the-eight-dimensions-of-quality

Hackman, J. R., & Wageman, R. (1995). Total quality management: Empirical, conceptual, and practical issues. *Administrative Science Quarterly*, *40*(2), 309–342. https://doi.org/10.2307/2393640

International Organization for Standardization. (2018a). *Food safety management systems: Requirements for any organization in the food chain* (ISO 22000:2018). https://www.iso.org/standard/65464.html

International Organization for Standardization. (2018b). *Quality management: Customer satisfaction: Guidelines for complaints handling in organizations* (ISO 10002:2018). https://www.iso.org/standard/71580.html

International Organization for Standardization. (2026a). *Quality management: Fundamentals and vocabulary* (ISO 9000:2026). https://www.iso.org/standard/9000

International Organization for Standardization. (2026b). *Quality management systems: Requirements* (ISO 9001:2026). https://committee.iso.org/standard/9001

International Organization for Standardization. (2026c). *Guidelines for auditing management systems* (ISO 19011:2026). https://www.iso.org/standard/19011

International Organization for Standardization. (2026d). *Environmental management systems: Requirements with guidance for use* (ISO 14001:2026). https://www.iso.org/standard/14001

International Organization for Standardization, Technical Committee 176. (2026, September 23). *ISO 9001:2026 has been released!* https://committee.iso.org/sites/tc176/home/news/content-left-area/news-and-updates/news.html

International Organization for Standardization/International Electrotechnical Commission. (2017). *General requirements for the competence of testing and calibration laboratories* (ISO/IEC 17025:2017). https://www.iso.org/standard/66912.html

Juran, J. M. (1988). *Juran on planning for quality*. Free Press.

Kaynak, H. (2003). The relationship between total quality management practices and their effects on firm performance. *Journal of Operations Management*, *21*(4), 405–435. https://doi.org/10.1016/S0272-6963(03)00004-4

Mahin, M., Kadasah, N., Alsabban, A., & Albliwi, S. (2024). Exploring the landscape of quality 4.0: A comprehensive review of its benefits, challenges, and critical success factors. *Production & Manufacturing Research*, *12*(1), Article 2373739. https://doi.org/10.1080/21693277.2024.2373739

McCollough, M. A., Berry, L. L., & Yadav, M. S. (2000). An empirical investigation of customer satisfaction after service failure and recovery. *Journal of Service Research*, *3*(2), 121–137. https://doi.org/10.1177/109467050032002

Montgomery, D. C. (2019). *Design and analysis of experiments* (10th ed.). Wiley. https://www.wiley.com/en-us/design-and-analysis-of-experiments-10th-edition-p-9781119492443

Montgomery, D. C. (2020). *Introduction to statistical quality control* (8th ed.). Wiley. https://www.wiley.com/en-us/introduction-to-statistical-quality-control-8th-edition-p-9781119723097

National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)*. https://www.nist.gov/itl/ai-risk-management-framework

National Institute of Statistics of Rwanda. (2026). *Integrated Business Enterprise Survey (2024)*. https://statistics.gov.rw/statistical-publications/business-establishment-finance-trade/business-establishment-finance-trade/integrated-business-enterprise-survey-2024

NIST/SEMATECH. (n.d.). *e-Handbook of statistical methods: Proportions control charts*. https://www.itl.nist.gov/div898/handbook/pmc/section3/pmc332.htm

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, *64*(1), 12–40.

Powell, T. C. (1995). Total quality management as competitive advantage: A review and empirical study. *Strategic Management Journal*, *16*(1), 15–37. https://doi.org/10.1002/smj.4250160105

Rwanda Standards Board. (2018). *Rwanda quality policy 2018*. https://www.rsb.gov.rw/fileadmin/user_upload/Quick_links/RWANDA_QUALITY_POLICY_2018.pdf

Rwanda Standards Board. (n.d.-a). *Overview of national standards*. https://www.rsb.gov.rw/standards/about-standards/overview

Rwanda Standards Board. (n.d.-b). *About national certification division*. https://www.rsb.gov.rw/certifications/about-ncd

Rwanda Standards Board. (n.d.-c). *Product certification*. https://www.rsb.gov.rw/certifications/product-certification

Samson, D., & Terziovski, M. (1999). The relationship between total quality management practices and operational performance. *Journal of Operations Management*, *17*(4), 393–409. https://doi.org/10.1016/S0272-6963(98)00046-1

Shah, R., & Ward, P. T. (2007). Defining and developing measures of lean production. *Journal of Operations Management*, *25*(4), 785–805. https://doi.org/10.1016/j.jom.2007.01.019

Shewhart, W. A. (1931). *Economic control of quality of manufactured product*. D. Van Nostrand.

Sila, I., & Ebrahimpour, M. (2005). Critical linkages among TQM factors and business results. *International Journal of Operations & Production Management*, *25*(11), 1123–1155. https://doi.org/10.1108/01443570510626925

Toyota Motor Corporation. (n.d.). *Basic concept of the Toyota Production System*. https://www.toyota-global.com/company/history_of_toyota/75years/data/automotive_business/production/system/change.html

Zu, X., Fredendall, L. D., & Douglas, T. J. (2008). The evolving theory of quality management: The role of Six Sigma. *Journal of Operations Management*, *26*(5), 630–650. https://doi.org/10.1016/j.jom.2008.02.001