# Making the Wrong Action Impossible: Software-Enforced Poka-Yoke in Automotive Assembly

## Architecture, Computational Methods, and Field Evidence from the MileSoft Pokayoke Suite

MileSoft Engineering Research Group
MileSoft Software Technologies
https://milesoft.net/
Corresponding authors: *Deepak Daryani* (deepakdaryani@milesoft.net), *Tanuj Daryani* (tanujdaryani@milesoft.net)

Working paper — Version 1.0 — 22 August 2026
Permanent link: https://milesoft.net/research/pokayoke

## Abstract

Shigeo Shingo drew a distinction that most quality programmes still blur: between a device that warns an operator of a deviation and a device that makes the deviating action impossible. Only the second is poka-yoke in his sense. The first is instrumentation, and its effectiveness is bounded by operator attention — which is exactly the variable that fails under time pressure, on night shifts, and on the sixth close-cousin variant of a part.

This paper presents the architecture, computational methods, and field evidence of the MileSoft manufacturing pokayoke suite: three station types (Pick-to-Light kitting, ADAS radar alignment, and end-of-line AC performance test) built on a shared station runtime in which the build recipe is governed master data, the interlock is the default control action, and every station outcome is bound to a unit identity at the moment it is produced.

We describe the control-versus-warning decision and why it is an architectural property rather than a configuration choice; the recipe model that lets six visually similar variants be distinguished without relying on operator discrimination; the measurement chain at a validation station and the guard-banding calculation that converts measurement uncertainty into acceptance limits; and the escape and false-reject accounting that makes a station's real performance visible.

A production deployment at a Tier-1 automotive supplier across plants at Pune and Nasik reports zero wrong-part picks since rollout against a pre-deployment operator error rate of approximately 1.4%, a 92% reduction in original-equipment-manufacturer containment incidents, operator ramp-up falling from four to six weeks to five days, and an 18% throughput gain on the flagship line. We close by proposing a capability reference framework for evaluating error-proofing stations.

**Keywords:** Poka-yoke; Error proofing; Source inspection; Pick-to-light; ADAS radar alignment; End-of-line test; Zero defect manufacturing; Measurement systems analysis; Guard banding; IATF 16949; ISO 26262; ISO 13849-1; First-pass yield; Automotive assembly

# 1 Introduction

Shingo's argument in Zero Quality Control is that inspection which detects defects after they occur cannot produce zero defects, however thorough it is, because it acts on the output rather than the cause. What produces zero defects is source inspection: detecting the condition that would cause a defect, at the moment it arises, and preventing the action that would realise it.

The distinction that follows is the one this paper is built on. A poka-yoke device has a setting function — how it detects a deviation — and a control function — what it does about one. Shingo separates control-type devices, which stop the process or make the wrong action impossible, from warning-type devices, which signal and leave the decision to the operator. Both are useful. Only the first is independent of attention.

Most industrial error-proofing deployments drift toward the warning type, for a reason that is organisational rather than technical: a warning never stops the line, and a control device sometimes does. The drift is rarely a decision anyone made; it accumulates through defaults. The architecture described here inverts the default so that the drift must be deliberate and recorded.

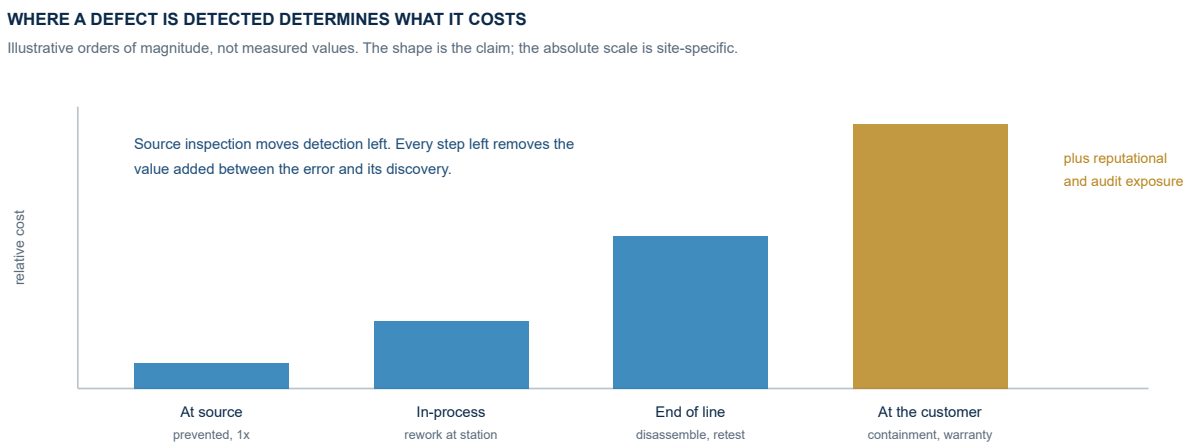


**Figure 1.** *Schematic.* Where a defect is detected determines what it costs. The bars are illustrative orders of magnitude rather than measured values; the claim is the shape, not the scale. Each step to the right adds the value that was built onto the defect before it was found, and the rightmost step adds containment and warranty exposure that is not proportional to the part's cost at all.

### 1.1 The variant discrimination problem

The deployment that motivated this work assembles radar modules in six variants that differ by small visual features and by substantially different downstream calibration. An operator distinguishing them by sight is performing a discrimination task under time pressure, and the observed error rate before deployment was approximately 1.4% — roughly one wrong pick in seventy.

That rate is not a training failure. It is close to what the human-factors literature would predict for a rapid discrimination task between similar items under production tempo. Training moves it somewhat; it does not move it to zero, and it does not hold under staff turnover. The only durable remedy is to remove the discrimination from the human task.

> If a process depends on an operator reliably telling two similar things apart at speed, the defect rate is a property of the process design, not of the operator.

### 1.2 Contributions

1. A station architecture in which the interlock is the default control action and warning-only operation requires explicit, recorded authorisation per joint or pick.
2. A recipe model that resolves the build specification from unit identity, removing variant discrimination from the operator's task.
3. An explicit measurement chain and guard-banding method for validation stations, converting measurement uncertainty into acceptance limits with the false-reject and escape trade stated.
4. Field evidence from a two-plant automotive deployment, and an eight-dimension capability reference framework distinguishing error-proofed stations from instrumented ones.

## 2 Background and Related Work

Three bodies of work bear on the design: the source-inspection tradition, the automotive quality and functional-safety regimes, and measurement systems analysis.

### 2.1 Source inspection and the poka-yoke taxonomy

Shingo classifies detection methods into contact, fixed-value, and motion-step types, and control responses into control and warning types. The classification is orthogonal: a contact sensor can drive either an interlock or a lamp. Which it drives is an engineering decision that the taxonomy makes visible and that most implementations leave implicit.

The related idea of successive and self-check inspection — that the next operation, or the operator themself, checks immediately rather than waiting for a downstream gate — is the organisational form of the same principle: shorten the interval between error and detection until it approaches zero, at which point detection becomes prevention.

## 2.2 Automotive quality and functional-safety regimes

IATF 16949 requires error-proofing devices to be verified, and requires a documented reaction plan for the case where a device fails. Both requirements have architectural consequences: verification must be scheduled and recorded per device, and the reaction plan must be enforceable rather than advisory, which in practice means the station must know what to do when its own sensor is suspect.

For radar modules specifically, the assembled product falls under ISO 26262, whose Part 7 addresses production, operation, service and decommissioning. The relevant consequence is that the alignment record produced at assembly is part of the safety argument for the finished vehicle, not merely a quality record, and its integrity requirements follow accordingly.

The interlocks themselves are safety-related parts of control systems in the sense of ISO 13849-1, which sets performance-level requirements for their design. A bin lock that fails open on power loss and a bin lock that fails closed are different designs with different performance levels, and the choice must be made against the hazard rather than convenience.

## 2.3 Measurement systems analysis

A validation station produces a verdict, and a verdict is only as trustworthy as the measurement chain behind it. ISO 5725 decomposes accuracy into trueness — closeness of the mean of many results to the reference value — and precision, with Part 2 giving the method for determining repeatability and reproducibility. The AIAG Measurement Systems Analysis manual provides the study designs that automotive practice uses for the same purpose.

The consequence too often skipped is that a station whose gauge variation is a significant fraction of the specification width will produce both false rejects and escapes at rates that no amount of tightening the limits can eliminate, because the limits are being applied to a noisy measurement. Section 4.3 states the guard-banding calculation that makes this explicit.

## 2.4 Incumbent practice and its failure modes

- Warning drift. Devices are commissioned as controls and quietly reconfigured to warnings after the first line stop, with no record that the change was made or why.
- Recipe on paper. The build specification lives on a printed sheet at the station; a revision reaches some stations and not others, and no record shows which revision built which unit.
- Unbound outcomes. A station records a pass or fail against a timestamp rather than a unit identity, so the record cannot answer a question about a specific vehicle later.
- Unexamined measurement. Acceptance limits are set equal to specification limits, with no allowance for gauge variation, so the escape rate is whatever the measurement noise makes it.

The first failure mode is the most damaging and the hardest to see, because a warning-configured station looks identical to a control-configured one on a plant walk.

# 3 System Overview and Architecture

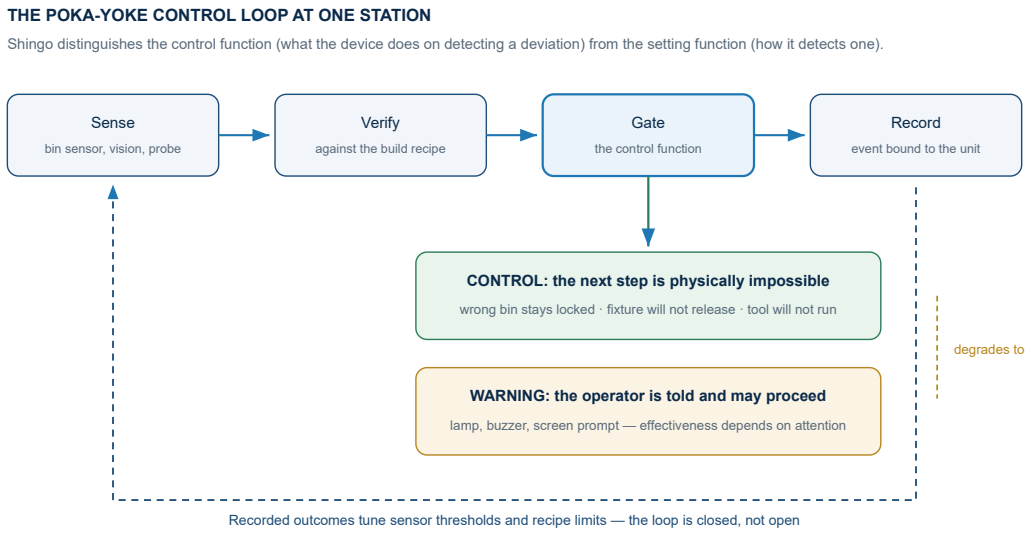


**Figure 2.** *Schematic.* The poka-yoke control loop at one station. The Gate step implements Shingo's control function. Its default outcome is the upper branch, in which the next step is physically impossible; the lower warning branch is available but must be authorised per joint or pick and is recorded as a deviation from the default. Recorded outcomes feed back into sensor thresholds and recipe limits, closing the loop. A station wired only to the lower branch is instrumented, not error-proofed — and on a plant walk the two look identical.

## 3.1 Design goals

- G1 — Control by default. An interlock is the default response to a deviation. Warning-only operation is configurable but requires explicit authorisation, recorded with an author and a reason.
- G2 — Recipe from identity. The build specification is resolved from the unit identity by the system, never selected by the operator.
- G3 — Bound outcomes. Every station outcome is written against a unit identity at the moment it is produced, not reconciled to one later.
- G4 — Versioned recipes. Recipes are versioned; every outcome records the version in force, so the specification that built a given unit is recoverable.
- G5 — Declared measurement. Every validation station declares its measurement chain and its combined uncertainty, and derives acceptance limits from them.
- G6 — Verified devices. Each error-proofing device carries a verification schedule; a device overdue for verification degrades the station to a defined reaction state rather than continuing silently.

### 3.2 The shared station runtime

All three station types run the same runtime, differing in their sensing hardware and their recipe schema. The runtime owns five things: identity resolution from a scan or fixture read; recipe resolution and version pinning; the gate decision and its interlock outputs; the outcome record; and the device verification state.

What the runtime deliberately does not own is the process logic of the operation itself. A radar alignment routine and an air-conditioning performance test have nothing in common computationally. Keeping the shared layer to identity, recipe, gate, record, and verification is what lets a third station type be added without touching the first two.

Station outcomes are streamed to a shared dashboard and to the plant's manufacturing execution system over OPC UA, positioning the runtime at operations level in the IEC 62264 sense with defined interfaces upward.

### 3.3 Degraded modes and the reaction plan

IATF 16949 requires a reaction plan for error-proofing device failure. The runtime implements this as an explicit station state rather than a document: a device that fails its verification check, reports an implausible reading, or exceeds its verification interval moves the station into a declared degraded state.

The degraded state is configured per station and is one of three: halt, in which the station refuses work; fallback, in which a secondary sensing path takes over at a stated lower confidence; or supervised, in which work continues under a named authoriser with every unit flagged for downstream re-check. What is not available is silent continuation, which is what an undeclared reaction plan amounts to in practice.

> A station that cannot state which of these three it is currently in does not have a reaction plan; it has a document describing one.

## 4 Computational Methods by Station Type

Notation is collected in Appendix A; worked numerical examples in Appendix B.

### 4.1 Pick-to-Light kitting

The station resolves the unit identity from a scan, resolves the pinned recipe version for that identity, and illuminates exactly the bin the recipe names. Every other bin is held locked. Confirmation is a sensor read at the bin, not a button press, so the record states that the correct bin was opened rather than that the operator asserted it was.

The effect on error rate is structural rather than statistical. Under a warning-only regime the residual wrong-pick probability is the product of the operator's discrimination error rate and the probability that the warning is not acted on. Under an interlock regime a wrong pick requires a device failure, and device failure is covered by the verification schedule and the reaction plan of Section 3.3.

$$p(\text{wrong} \mid \text{warning}) = p(\text{discrimination error}) \times (1 - p(\text{warning acted on})) \quad \text{versus} \quad p(\text{wrong} \mid \text{control}) = p(\text{device failure})$$ (pickerr)

where p(d) is the operator discrimination error rate, p(a) the probability the warning is noticed and acted on, and p(f) the interlock failure probability. The two regimes differ in kind: the first scales with human factors, the second with device reliability and verification interval.

Ramp-up time follows from the same shift. When variant discrimination is the system's job rather than the operator's, a new operator must learn the physical motion and the station's rhythm, not the differences between six similar parts. Section 5 reports the observed effect.

## 4.2 ADAS radar alignment

The alignment station seats the module in a fixture, reads a fiducial by vision to establish the datum, drives a sealed alignment routine against a calibrated reference target, and emits a signed alignment report bound to the module identity. The routine is sealed in the sense that its parameters come from the pinned recipe version and cannot be adjusted at the station.

The verdict rests on the measurement chain, and the chain must be declared (G5). Its links are the calibrated reference artefact, the fixture's datum repeatability, the sensor's bias and resolution, and the method's sampling and filtering. Each contributes a variance component, and independent contributions combine in the usual way.

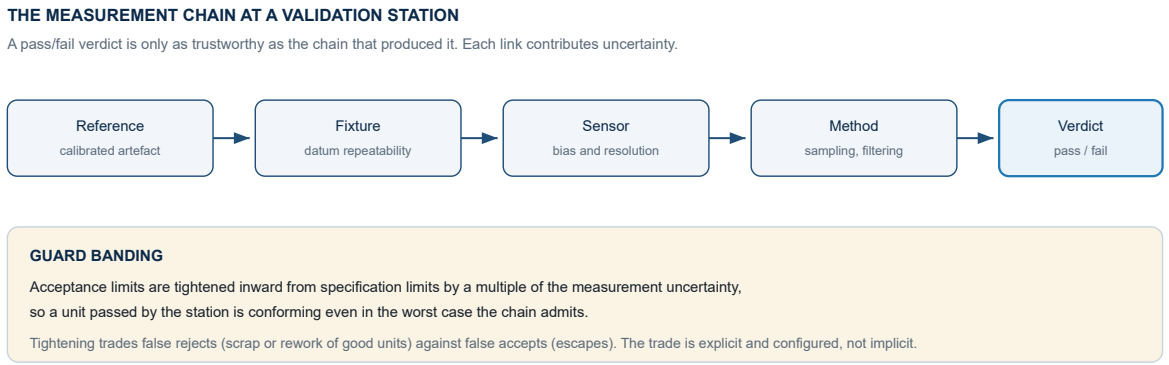


**Figure 3.** *Schematic.* The measurement chain at a validation station, and the guard band that follows from it. Acceptance limits are tightened inward from specification limits by a multiple of the combined uncertainty, so that a unit the station passes is conforming even in the worst case the chain admits.

$$u(\text{combined}) = \text{sqrt}( u(\text{reference})^2 + u(\text{fixture})^2 + u(\text{sensor})^2 + u(\text{method})^2 )$$ (uc)

the combined standard uncertainty of the alignment measurement, assuming the four contributions are independent. Where a fixture and a sensor share a thermal drift they are not independent, and a correlation term must be added rather than assumed away.

Because the assembled module falls under ISO 26262, the alignment record forms part of the finished vehicle's safety argument. The record therefore carries the recipe version, the reference artefact's calibration state, and the combined uncertainty in force at the time, not merely the measured angle and a verdict.

## 4.3 Guard banding and the accept/reject trade

Setting acceptance limits equal to specification limits guarantees escapes at a rate determined by measurement noise, because a unit measured just inside the limit may be truly outside it. Guard banding tightens the acceptance limits inward.

$$AL(upper) = USL - k \times u(combined),\ AL(lower) = LSL + k \times u(combined) \quad \text{(guard)}$$

where k is the guard-band multiplier, chosen from the tolerable escape probability. k = 2 corresponds to roughly 95% confidence for a normally distributed measurement error; safety-relevant characteristics are typically banded harder.

The trade is explicit. Increasing k reduces escapes and increases false rejects, which are good units scrapped or reworked. Both costs are real, and the correct k depends on their ratio, which differs between a safety-relevant alignment and a cosmetic characteristic. The architecture requires k to be configured per characteristic rather than defaulted globally, precisely so the trade is made rather than inherited.

The measurement system's own adequacy is assessed by gauge repeatability and reproducibility against the tolerance width, following the AIAG study designs.

$$\%GRR = 100 \times \sigma(GRR) \,/\, (USL - LSL) \quad \text{(grr)}$$

expressed against tolerance rather than against total variation, because the question at a validation station is whether the gauge can resolve conformance, not whether it can resolve part-to-part variation. Automotive practice treats below 10% as acceptable and above 30% as unacceptable.

## 4.4 End-of-line AC performance test

The air-conditioning performance station is a functional test rather than a dimensional one: the unit is run against a defined duty and its measured performance compared against the recipe's envelope. The same runtime applies — identity, pinned recipe, gate, bound record, device verification — with the sensing being a set of temperature, pressure, and current measurements rather than a vision read.

Two properties distinguish functional testing from dimensional testing, and both affect the gate. Measurements are time-dependent, so the recipe specifies a settling period before the acceptance window is evaluated; and the test consumes cycle time, so the station is frequently the line's constraint. The runtime records settling time separately from measurement time, which makes it possible to distinguish a slow test from a slow product.

### 4.5 Yield and escape accounting

Station performance is reported as first-pass yield, and line performance as the rolled product of station yields — the measure that reveals how a line of individually respectable stations produces a poor finished-unit yield.

$$RTY = FPY(1) \times FPY(2) \times \ldots \times FPY(n) \quad \text{(rty)}$$

where FPY(i) is the first-pass yield of station i. Ten stations at 99% each give RTY = 0.904, so roughly one unit in ten requires intervention somewhere — a result no single station's report would suggest.

Escapes are counted separately from failures, because they are found downstream and are the number that predicts containment exposure. A station's escape count is knowable only from downstream detection, so it is attributed retrospectively to the station whose recipe should have caught the condition, which is possible only because outcomes are bound to unit identities (G3).

## 5 Field Evidence: A Production Deployment

The deployment described here is at a Tier-1 automotive supplier operating plants at Pune and Nasik, India, on a radar-module assembly line. All figures are operator-reported from the production system and are itemised with their provenance below.

### 5.1 Context and prior workflow

Radar modules ship in six close-cousin variants with small visual differences and substantially different downstream calibration. The operator pick-error rate hovered around 1.4%. Off-specification radar alignment passed quality control and surfaced at the original equipment manufacturer, generating containment incidents and damaging the customer relationship.

Training a new operator to hold cycle time on more than two variants took four to six weeks. Audit preparation for customer line walks consumed a senior process engineer's full week each quarter.

### 5.2 What was deployed

Pick-to-Light was installed on every kitting bay with bins physically locked until the correct pick is confirmed. The ADAS radar alignment station was integrated with the manufacturing execution system so each module boots through a sealed alignment routine and ships a signed alignment report. An AC performance test station was added downstream as a cross-check on a sister product family running on the same line. All three stream their events to one dashboard for shift leads and audit playback.

### 5.3 Reported outcomes

**Table 1.** Operator-reported outcomes following deployment, measured against the pre-deployment baseline on the same line.

| Measure | Before | After |
|---|---|---|
| Operator wrong-part picks | approx. 1.4% error rate | 0 since rollout |
| OEM containment incidents | Baseline | 92% reduction |
| Operator ramp-up to full cycle time | 4-6 weeks | 5 days |
| Throughput per shift, flagship line | Baseline | +18% |
| Variants handled without operator discrimination | 0 of 6 | 6 of 6 |

The zero result should be read precisely. It is zero wrong-part picks recorded since rollout, under a regime in which a wrong pick requires an interlock failure rather than an operator error. It is not a claim that the failure probability is zero; it is a claim that the failure mode changed from one governed by human factors to one governed by device reliability, and that no device failure has produced a wrong pick within the observation window.

The throughput gain is the least intuitive result and the most instructive. Interlocks were expected to cost cycle time, because a locked bin is a pause. In practice they removed more time than they added: operators stopped double-checking, and the rework loop for wrong picks — which had consumed both the station's time and the downstream station's — largely disappeared.

The ramp-up reduction follows directly from Section 4.1. Five days is the time to learn a motion and a rhythm; four to six weeks was the time to learn to tell six similar parts apart reliably at speed. Removing the second task from the operator removed it from the training curve.

## 6 Discussion

### 6.1 The default is the design

Of the six design goals, G1 carries most of the result. The technical capability to interlock a bin is not novel and was available in the incumbent tooling. What was absent was a default that made interlocking the normal case and warning-only an authorised exception.

This matters because the drift toward warnings is driven by a real and legitimate pressure. A control device stops the line, and stopping the line is visible, costly, and attributable, while an escape is invisible until it reaches the customer. Any architecture that leaves the choice to local discretion under that asymmetry will converge on warnings. Making the exception explicit and recorded does not remove the pressure; it makes yielding to it visible.

## 6.2 Guard banding as an honesty requirement

A station whose acceptance limits equal its specification limits is passing units it cannot distinguish from failing ones. This is not a subtle statistical point; it is arithmetic that follows from any non-zero measurement uncertainty. Yet acceptance-equals-specification remains a common default, because guard banding visibly increases the reject rate and a higher reject rate looks like worse performance.

The reject rate does rise, and this paper does not present that as costless. What changes is which error the station makes: it moves from silently passing marginal units to visibly rejecting some good ones. For a safety-relevant characteristic under ISO 26262 that is the correct direction; for a cosmetic one it may not be. Requiring k to be configured per characteristic is what forces the question to be asked.

## 6.3 A capability reference framework for error-proofing stations

**Table 2.** Capability reference framework. Each dimension distinguishes a genuinely error-proofed station from an instrumented one, and each is answerable by demonstration at the station within minutes.

| Dimension | Question the station must answer by demonstration |
|---|---|
| D1 Control by default | Attempt the wrong action. Is it prevented, or merely announced? |
| D2 Authorised exceptions | Can the station list every characteristic in warning-only mode, with who authorised it and why? |
| D3 Recipe from identity | Does the operator ever select the variant, or does the system resolve it from the unit? |
| D4 Version recovery | For a unit built six months ago, can the station state which recipe version built it? |
| D5 Declared measurement | Can the station state its measurement chain and its combined uncertainty? |
| D6 Guard band | Are acceptance limits tighter than specification limits, and is k configured per characteristic? |
| D7 Device verification | What happens when a device is overdue for verification? Demonstrate the degraded state. |
| D8 Escape attribution | Can a downstream escape be attributed back to the station whose recipe should have caught it? |

## 6.4 Generalisability

The evidence base is one operator, two plants, one product family. The variant-discrimination result should generalise wherever similar-looking parts are picked under tempo, which is common across automotive and appliance assembly. The alignment and guard-banding results are specific to stations producing a measured verdict, and their magnitude depends on the ratio of measurement uncertainty to tolerance width — a ratio that varies widely and must be established per station rather than assumed.

## 7 Threats to Validity and Limitations

1. Single-operator evidence. All field figures come from one supplier, two plants, and one product family. There is no control line and no matched comparison against an alternative error-proofing approach.
2. A zero is a censored observation. Zero wrong picks since rollout bounds the rate below the reciprocal of the units built, but it does not establish that the rate is zero. The observation window and unit count required to make that bound meaningful are not published here.
3. Operator-reported metrics. Figures are reported from the production system by the operator rather than independently audited.
4. Confounded deployment. Three station types were introduced alongside process and training change; no single station's contribution to the throughput or containment figures can be isolated.
5. Throughput mechanism inferred. The explanation offered in Section 5.3 for the throughput gain — reduced double-checking and rework — is consistent with the operator's account but was not measured directly by time study.
6. No cost data. Operational outcomes are reported; device, integration, and maintenance costs are not, so no return-on-investment claim is made or implied.

> The second limitation deserves emphasis because zero is the most rhetorically powerful and least statistically informative result a paper of this kind can report.

## 8 Future Work

- Bounding the zero. Publishing units built and the observation window alongside the zero count would convert a rhetorical result into a statistical bound on the interlock failure rate.
- Direct measurement of the throughput mechanism. A time study separating double-check time, rework time, and interlock wait time would test the explanation offered in Section 5.3 rather than inferring it.
- Guard-band optimisation from realised costs. Choosing k per characteristic currently rests on judgement; deriving it from observed false-reject and escape costs would make the trade quantitative.
- Warning-mode telemetry. Recording how often a warning is issued and not acted upon would give an empirical value for the term p(a) in Equation (pickerr), which is currently an assumption.
- Cross-plant device-reliability data. Aggregating interlock failure and verification-overdue events across sites would let p(f) be estimated rather than bounded by the absence of observed failures.

## 9 Conclusion

Shingo's distinction between control and warning devices is sixty years old and still the most consequential decision in an error-proofing programme. The finding of this work is that the decision is rarely made explicitly: it is inherited from defaults, and the defaults drift toward warnings because stopping a line is

visible while an escape is not.

This paper has described a station architecture that inverts the default — interlock as the normal case, warning-only as an authorised and recorded exception — together with a recipe model that removes variant discrimination from the operator's task, and an explicit measurement chain with guard banding that converts measurement uncertainty into acceptance limits. A two-plant production deployment reports zero wrong-part picks since rollout against a 1.4% baseline error rate, a 92% reduction in customer containment incidents, operator ramp-up falling from four to six weeks to five days, and an 18% throughput gain.

The capability reference framework of Section 6.3 is offered as the durable contribution, and its first question is the one that matters most: attempt the wrong action, and observe whether the station prevents it or merely announces it.

---

# Appendix A Nomenclature

**Table 3.** Symbols and abbreviations used in this paper.

| Symbol / term | Meaning |
|---|---|
| p(d) | Operator discrimination error rate at a picking task |
| p(a) | Probability that an issued warning is noticed and acted upon |
| p(f) | Interlock device failure probability |
| u(ref), u(fix), u(sen), u(met) | Uncertainty contributions from reference, fixture, sensor, and method |
| u(combined) | Combined standard uncertainty of the measurement chain |
| USL, LSL | Upper and lower specification limits for a characteristic |
| AL(upper), AL(lower) | Guard-banded acceptance limits applied by the station |
| k | Guard-band multiplier, configured per characteristic |
| sigma(GRR) | Standard deviation attributable to gauge repeatability and reproducibility |
| %GRR | Gauge repeatability and reproducibility as a percentage of tolerance width |
| FPY | First-pass yield of a single station |
| RTY | Rolled throughput yield, the product of station first-pass yields |
| Escape | A non-conforming unit passed by a station and detected downstream |
| False reject | A conforming unit rejected by a station |
| Control device | A poka-yoke device that makes the wrong action impossible |
| Warning device | A poka-yoke device that signals a deviation and leaves the decision to the operator |
| ADAS | Advanced driver-assistance system |
| MES | Manufacturing execution system |

# Appendix B Worked Numerical Examples

## Appendix B.1 Expected wrong picks under warning versus control

A line builds 690 units per shift, each requiring four picks from close-cousin bins, so 2,760 picks per shift. The observed operator discrimination error rate before deployment was p(d) = 0.014.

Under a warning-only regime, suppose a warning is noticed and acted upon 80% of the time, so p(a) = 0.8. Applying Equation (pickerr): p(wrong) = 0.014 x (1 — 0.8) = 0.0028, giving 2,760 x 0.0028 = 7.7 wrong picks per shift.

Under an interlock regime the wrong-pick probability is the device failure probability. For an interlock verified daily with a demonstrated failure rate on the order of 1e-5 per actuation, the expected count is 2,760 x 1e-5 = 0.028 per shift, or roughly one per thirty-six shifts.

The comparison is instructive precisely because p(a) is an assumption. Even at a generous p(a) = 0.95, the warning regime yields 1.9 wrong picks per shift — still two orders of magnitude worse than the interlock. The conclusion does not depend on the exact value, which is why removing the term matters more than estimating it well.

## Appendix B.2 Guard banding an alignment characteristic

A radar alignment characteristic is specified at 0.00 degrees with limits of plus or minus 0.30 degrees, so USL = 0.30 and LSL = -0.30. The declared measurement chain contributes u(reference) = 0.020, u(fixture) = 0.045, u(sensor) = 0.030, and u(method) = 0.015 degrees.

Applying Equation (uc): u(combined) = sqrt(0.020^2 + 0.045^2 + 0.030^2 + 0.015^2) = sqrt(0.000400 + 0.002025 + 0.000900 + 0.000225) = sqrt(0.003550) = 0.0596 degrees.

With k = 2, Equation (guard) gives AL(upper) = 0.30 — 2 x 0.0596 = 0.181 degrees and AL(lower) = -0.181 degrees. The station accepts a band 60% as wide as the specification. A unit measuring 0.25 degrees — comfortably inside specification — is rejected, because the chain cannot distinguish it from a unit truly at 0.37 degrees.

The fixture dominates the uncertainty budget at 0.045 of 0.0596 combined. Halving fixture repeatability would give u(combined) = 0.0424 and widen the acceptance band to plus or minus 0.215 degrees. That is the correct engineering response: improve the largest contributor rather than reduce k and accept more escapes.

## Appendix B.3 Is the gauge adequate?

For the same characteristic, a gauge study returns sigma(GRR) = 0.052 degrees against a tolerance width of USL — LSL = 0.60 degrees.

Applying Equation (grr): %GRR = 100 x 0.052 / 0.60 = 8.7%. This is below the 10% threshold automotive practice treats as acceptable, so the measurement system is adequate to judge conformance against this tolerance.

Had the tolerance been plus or minus 0.10 degrees instead, the width would be 0.20 and %GRR would be 26% — in the marginal band, and adequate only with justification. The same gauge is adequate or inadequate depending entirely on the tolerance it is asked to judge, which is why %GRR is expressed

against tolerance rather than against part variation at a validation station.

### Appendix B.4 Why respectable stations make a poor line

A line has twelve stations. Eleven run at 99.5% first-pass yield and one — the AC performance test — runs at 96.0%.

Applying Equation (rty): RTY = 0.995^11 x 0.960 = 0.9464 x 0.960 = 0.9086. Roughly one finished unit in eleven required intervention somewhere on the line.

Improving the weakest station from 96.0% to 99.0% raises RTY to 0.9464 x 0.990 = 0.9369. Improving all eleven others from 99.5% to 99.8% instead raises it to 0.9782 x 0.960 = 0.9391. The two interventions are comparable in effect, which is not what a station-by-station report suggests — and only the rolled measure reveals it.

## Data provenance

Every quantitative claim in this paper resolves to one of the entries below. Figures marked as modelled estimates are not deployment measurements; their assumptions are stated.

| Metric | Reported value | Provenance |
|---|---|---|
| Wrong-part picks since rollout | 0 | **Operator-reported — Tier-1 Manufacturers (Pune & Nasik plant, India)** /case-studies/pokayoke — Baseline operator pick-error rate was approximately 1.4%. |
| Pre-deployment operator pick-error rate | approx. 1.4% | **Operator-reported — Tier-1 Manufacturers (Pune & Nasik plant, India)** /case-studies/pokayoke — Six close-cousin radar-module variants with small visual differences. |
| OEM containment incidents | 92% reduction | **Operator-reported — Tier-1 Manufacturers (Pune & Nasik plant, India)** /case-studies/pokayoke |
| Operator ramp-up to full cycle-time competence | 5 days (from 4-6 weeks) | **Operator-reported — Tier-1 Manufacturers (Pune & Nasik plant, India)** /case-studies/pokayoke |
| Throughput per shift, flagship line | +18% | **Operator-reported — Tier-1 Manufacturers (Pune & Nasik plant, India)** /case-studies/pokayoke |
| Wrong-part reduction attributed to Pick-to-Light | over 90% | **Product specification** /products/pick-to-light |

## References

**[1]** Shingo, S. (translated by A. P. Dillon) (1986). Zero Quality Control: Source Inspection and the Poka-Yoke System. Productivity Press, Cambridge, MA. Originally published in Japanese, 1985.

**[2]** Nakajima, S. (1988). Introduction to TPM: Total Productive Maintenance. Productivity Press, Cambridge, MA.

**[3]** International Automotive Task Force (2016). Quality management system requirements for automotive production and relevant service parts organizations. IATF 16949:2016, first edition, 1 October 2016 (superseding ISO/TS 16949).

**[4]** International Organization for Standardization (2018). Road vehicles - Functional safety - Part 1: Vocabulary. ISO 26262-1:2018. https://www.iso.org/standard/68383.html

**[5]** International Organization for Standardization (2018). Road vehicles - Functional safety - Part 2: Management of functional safety. ISO 26262-2:2018. https://www.iso.org/standard/68384.html

**[6]** International Organization for Standardization (2018). Road vehicles - Functional safety - Part 4: Product development at the system level. ISO 26262-4:2018.

**[7]** International Organization for Standardization (2018). Road vehicles - Functional safety - Part 7: Production, operation, service and decommissioning. ISO 26262-7:2018.

**[8]** International Organization for Standardization (2015). Safety of machinery - Safety-related parts of control systems - Part 1: General principles for design. ISO 13849-1:2015. https://www.iso.org/standard/69883.html

**[9]** International Organization for Standardization (2023). Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions. ISO 5725-1:2023 (superseding ISO 5725-1:1994). https://www.iso.org/standard/69418.html

**[10]** International Organization for Standardization (2019). Accuracy (trueness and precision) of measurement methods and results - Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method. ISO 5725-2:2019. https://www.iso.org/standard/69419.html

**[11]** Automotive Industry Action Group (Chrysler, Ford, General Motors) (2010). Measurement Systems Analysis (MSA) Reference Manual. AIAG, 4th edition, MSA-4.

**[12]** International Organization for Standardization (2014). Statistical methods in process management - Capability and performance - Part 1: General principles and concepts. ISO 22514-1:2014. https://www.iso.org/standard/64135.html

**[13]** International Organization for Standardization (2017). Statistical methods in process management - Capability and performance - Part 2: Process capability and performance of time-dependent process models. ISO 22514-2:2017. https://www.iso.org/standard/71617.html

**[14]** Verein Deutscher Ingenieure / Verband der Elektrotechnik (2025). Minimum restrictions for application of fastening systems and tools - Applications in the automotive industry. VDI/VDE 2862 Blatt 1.

**[15]** International Organization for Standardization (2017). Assembly tools for screws and nuts - Hand torque tools - Part 1: Requirements and methods for design conformance testing and quality conformance testing: minimum requirements for declaration of conformance. ISO 6789-1:2017. https://www.iso.org/standard/62549.html

**[16]** International Organization for Standardization (2017). Assembly tools for screws and nuts - Hand torque tools - Part 2: Requirements for calibration and determination of measurement uncertainty. ISO 6789-2:2017. https://www.iso.org/standard/62550.html

**[17]** International Organization for Standardization (2014). Automation systems and integration - Key performance indicators (KPIs) for manufacturing operations management - Part 2: Definitions and descriptions. ISO 22400-2:2014. https://www.iso.org/standard/54497.html

**[18]** International Electrotechnical Commission / International Society of Automation (2025). Enterprise-control system integration - Part 1: Models and terminology. IEC 62264-1; ANSI/ISA-95.00.01-2025 (IEC 62264-1 Mod).

**[19]** International Electrotechnical Commission (2025). OPC unified architecture - Part 1: Overview and concepts. IEC 62541-1:2025. https://webstore.iec.ch/en/publication/81513

**[20]** Lee, J., Bagheri, B., and Kao, H.-A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manufacturing Letters, 3, 18-23.