



Designing & Creating Test Procedures

Reference Guide

How to drive accuracy, consistency, completeness, & defensibility

How to use this deck

- This is a reference guide, not a “read straight through” training
- Start with Section 1 to select the right OE check intent
- Use Section 2 to write deterministic Pass/Fail logic
- Use Section 3 to choose evidence format that can actually prove the claim
- If evidence can't prove it, change the evidence requirement or narrow scope

Outcomes

- Procedures become mechanically decidable (binary, repeatable)
- Evidence requirements are aligned to what needs to be proven
- Correlation/completeness tests move from screenshots to text-based exports
- Fewer “random fails” caused by truncation, UI wording variance, and time anchoring errors

Understanding what OE checks are designed TO DO & NOT DO

What an OE check is

- A structured, repeatable method to determine Pass/Fail from evidence
- Focused on proving a specific claim (not the entire control narrative)
- Designed to be defensible even when reviewed by someone who wasn't involved

What an OE check is NOT

- Not a "review for reasonableness"
- Not a catch-all validation of process maturity
- Not a narrative summary disguised as a test
- Not a place to infer intent when evidence is ambiguous

The core contract

- If the evidence shown meets the criteria → Pass
- If any required element is missing/unclear/incomplete → Fail (or "insufficient evidence" depending on your org's taxonomy)
- The procedure must state:
 - exactly what must be visible
 - exactly how to interpret it
 - what makes evidence unusable

When to split an OE check

- If one procedure tries to prove: configuration + monitoring + alerting + reviews + patching
- Split into smaller checks:
 - Configuration check
 - Operational evidence check (alerts/reviews)
 - Completeness check (population coverage)
 - Frequency check (occurrences)

Common mismatch: what people want vs what evidence can prove

Want	Evidence Needed
"only approved people can edit"	full access list export + approved roster (join key)
"priority is system-calculated and locked"	screenshots prove recalculation; enforcement requires config proof
"quarterly meetings occurred"	an index/listing of occurrences (not just one set of minutes)

The 3 outcomes an OE check can legitimately prove

- Configuration/value is set as required (point-in-time)
- Behavior occurs as expected when inputs change (within-object behavior)
- Population rule holds across a dataset (completeness/correlation)

Design of Test Procedures

Golden rules for consistent procedures

- Use deterministic verbs: confirm, match, equals, is visible, is set to, corresponds to
- Avoid interpretive verbs: validate, ensure, appropriate, adequate, reasonable
- Prefer explicit conditions over narrative explanations
- Use one source-of-truth system per test (unless correlation is the point)

Decision flow (pick the right OE Check Type)

- Are you checking one value in one place? → A
- Multiple settings in one config view? → B
- Proving recalculation/behavior in the same object? → C
- "All users/domains/rows must..."? → D
- Matching roster/list A to list B? → E
- Any "before/after" date logic? → F
- "Monthly/quarterly" expectation? → G

Standard test procedure format [Short title]: Review the [artifact] and verify [specific condition]. Fail if [condition is not met].

- Test name (short, specific & describes single atomic thing being tested)
- Allowed evidence (system + exact view/export)
- Inputs/fields evaluated (specific list)
- Acceptable equivalents (if UI wording varies)
- Truncation/completeness rule (what makes evidence unusable)
- Fail criteria

OE check types (quick taxonomy)

- Type A: Single-field check (one artifact)
- Type B: Multi-field config check (one artifact)
- Type C: Within-record behavior check (same record, multiple screenshots)
- Type D: Population completeness check (tabular required)
- Type E: Correlation check (two sources + join key; tabular required)
- Type F: Time-anchored rule check (tabular preferred; anchor defined)
- Type G: Frequency/occurrence check (index required)

Quality of Evidence used — based on test type

Evidence format rules by OE check type

- Type A/B/C → screenshots allowed if readable and complete
- Type D/E/F/G → text-based tabular evidence required (CSV/XLSX/export/text list)
- Hard line:
 - screenshots are for checking items visible on the screenshot
 - screenshots are not acceptable for correlation/completeness

Correlation requirements (when matching across sources)

- Must define join key:
 - email/userID > display name
 - ticket number/sys_id > subject line
- Must define normalization if names are used:
 - case-insensitive match
 - diacritics handling (if applicable)
- If join key isn't available, require a better evidence source

"Insufficient evidence" patterns

- Scrollable access lists
- Composite screenshots missing the key rows
- Screenshots where the value field is cut off
- "Diagnostic" pages used when the procedure requires a specific system-of-record table
- Reports where the comparison anchor is unclear

Screenshot quality gates (must meet)

- Readable text (no blur, no tiny illegible fields)
- Field label + field value visible together
- System context visible when system-of-record matters (header/breadcrumb/recognizable UI)
- No truncation indicators if completeness is required
- If chips/list overflow exists, screenshots are insufficient for inclusion checks

Tabular evidence quality gates (must meet)

- Includes stable identifiers (email/userID/assetID preferred over names)
- Includes all required columns (status, dates, permissions)
- Anchor date is present or clearly tied to the export timestamp
- Export covers full population required by the procedure

Evidence bundling (how to submit cleanly)

- Name evidence files predictably: System_View_ObjectID_Date
- For multi-screenshot behavior tests: include the same object ID in every screenshot
- For exports:
 - include export/run timestamp if possible
 - avoid partial extracts unless procedure explicitly allows sampling

Appendices

Test Procedure Check (inconsistency triggers list)

- Vague verbs
- Missing anchor definition
- Screenshot used for correlation
- Exact string match where UI varies
- No truncation rule
- AND/OR logic ambiguous
- Over-scoped (tries to prove too much)

QA checklist (publish gate)

- Deterministic Pass/Fail
- Evidence type matches OE type
- Truncation rule defined
- Equivalents defined
- Time anchor defined (if applicable)
- Join key defined (if applicable)