

Validation plan

Your goal: Design a test plan that defines the metrics needed to validate your prototype.

One product: Validation plan with prototype claim, two measurable metrics with units and pass/fail thresholds, a control, a test procedure, and one error-reduction step.

Question: What exactly will you check, what counts as passing, and what will that result still leave unknown?

Start the work

Multiply each anchored score by its weight, then explain the tradeoff and the limits of the resulting rank.

1. State the specific claim your prototype is supposed to satisfy.
2. Choose two measurable validation metrics with units.
3. Describe the test procedure, including what counts as a pass or fail.
4. Identify a control or baseline for comparison.
5. List one source of error and how you would reduce it.

Save the same record locally. List the missing assignment, authentic design, sources or protocol without claiming completion.

Words you will use

prototype

An early representation of a proposed design.

validation plan

A plan for evaluating intended use with stated evidence and limits.

verification

Checking whether a product meets specified requirements.

metric

A defined quantity used to assess a stated claim.

unit

The named quantity being counted or measured.

threshold

A declared boundary for a decision.

pass/fail rule

A rule that classifies a result against a requirement.

control / baseline

A reference condition interpreted alongside the design being checked.

procedure

Ordered instructions another person can follow.

repeat

A new inspection or run using the stated procedure.

sample

One named item in the check, not necessarily a person.

minimum

The smallest listed value.

qualifying field

Here, a distinct Q label with an adjacent blank marker.

instruction type

One unique required instruction in the S1 checklist.

version

An identified form of a design or record.

error

A mistake or measurement problem that can distort a result.

error reduction

A step intended to reduce a specified error.

reproducibility

The ability to follow documented methods and obtain consistent conclusions in the stated context.

confounder

A varying factor that complicates attributing a result to the intended difference.

causation

One factor producing a change in another; a comparison alone does not establish it.

blind review

A review that hides relevant identities or conditions from an assessor.

sham

A comparison that mimics an intervention while omitting its active element.

expectation bias

A judgment influenced by the result someone expects.

usability

How intended users interact with a product in context.

structural claim

A claim about specified parts or features, not user outcomes.

simulated reading

A clearly invented input used to practice analysis.

limitation

A boundary on what evidence can establish.

pre-lab

A written plan made before an actual test.

stakeholder

A person or group affected by a design.

millimeter (mm)

A length unit equal to one thousandth of a meter.

E1: match the check to the claim

NASA distinguishes verification against specified requirements from validation against intended use and stakeholder expectations. A plan is not a completed test. Document unresolved failures and limits before claiming completion.

NIST recommends writing and prioritizing experimental objectives before choosing factors, responses and a design. State what the test can answer before selecting a number to measure.

Today's recovery checks structural requirements of paper records as one early part of a validation plan. It does not validate how people use the aid. A control and a threshold improve interpretation; they do not alone prove causation.

E2: paper completeness is not health or usability

FDA describes human factors through how people perceive, interpret and act on device information. Labels and training materials are part of the interface. Safe and effective use needs evidence about use, not merely the presence of text.

Counting question fields cannot establish comfort, understanding or better health. A lower heart-rate number also does not by itself mean a person is healthier. This exercise prescribes no clinical threshold or intervention.

Fictional caution: an invented team says a bandage helps wounds heal faster but supplies no comparison or defined outcome. The claim cannot be assessed from that statement. No such team or result is asserted to be real.

S1: adopted design B and baseline A

Use your actual March 25 selection and evidence when available. For complete original recovery, explicitly adopt B from the supplied paper plans. This does not claim your matrix selected B, that you built it, or that it has been used.

B specifies three sheets, three assembly actions, five blank question fields and four instruction types. These counts match B's March 25 paper plan. Its narrow claim is that a reproduced set retains all five fields and four instructions.

Complete B text plan: sheet 1 has Q1 [blank] and Q2 [blank]; sheet 2 has Q3 [blank] and Q4 [blank]; sheet 3 has Q5 [blank]. The bracketed markers stand for blank response regions in this text representation, not a measured print size.

The four required instruction types are: Write a question; Circle a priority; Bring the sheet; Leave a box blank if unsure. All four appear on sheet 1. No question is filled in.

Baseline A specifies one sheet and one assembly action. Its complete text plan is Q1 [blank], Q2 [blank], Write a question, Circle a priority. It has fewer planned fields and instructions, not a proven difference in usefulness.

These text records support counting only. Actual writable area, print legibility, materials, audio and user understanding remain untested. Use no names, personal health questions, patient records or human intervention.

R1: two metrics, units and fixed pass rules

Metric 1 is the number of distinct Q labels with an adjacent blank marker per copy, in fields/copy. Count a repeated Q label only once. A missing blank marker does not qualify.

For B, each of three copies must have at least 5 qualifying fields. Report the minimum count across the three copies. Pass if that minimum is at least 5 fields/copy; otherwise fail.

Metric 2 is the number of the four required instruction types present per copy, in types/copy. Use S1's exact four-type checklist. Count each type once; extra repeated wording does not add coverage.

Each of three copies must contain all 4 types. Report the minimum count. Pass if that minimum is 4 types/copy; otherwise fail. The set passes this narrow screen only if both metrics pass.

These thresholds are declared original planning requirements, not medical standards, surveyed needs or requirements from NASA. Set them before classifying samples. Do not lower a threshold after seeing a failure.

Inspect A with the same definitions and record its counts as a comparison baseline. A comparison difference cannot isolate the effects of extra sheets, fields or wording.

R2: repeatable procedure and blank test-plan structure

Write a plan with these labeled entries: version and source; structural claim; two metrics and units; thresholds; baseline; materials; ordered procedure; three sample IDs; error and fix; result rule; limitation.

For original planning, use the full S1 text records, a copy of the four-type checklist, paper and pencil or a local document. Describe three reproduced copies per design. Do not claim actual materials or copies exist if they do not.

Procedure 1: freeze the version, S1 definitions and both R1 thresholds. Give copies IDs B1-B3 and A1-A3. For a future paper check, reproduce each copy from its plan without adding questions.

Procedure 2: inspect each copy independently. List every distinct qualifying Q label, count the fields, then mark yes or no for each required instruction. Record the two counts before looking at earlier totals.

Procedure 3: inspect again in reverse copy order using the same checklist. If readings disagree, retain both readings, inspect the exact text and document a corrected count with a reason.

Procedure 4: compute each design's minimum field count and minimum instruction count. Apply each predeclared threshold. Record pass or fail for both metrics and the combined result; keep all failed copies visible.

Likely error: counting a duplicate label twice or missing an instruction. Fix: use a distinct-label list and the fixed four-item checklist on both passes. This reduces counting error but does not test usability.

Blank result log columns: design/version; copy ID; qualifying Q labels; field count; four yes/no instruction entries; type count; field pass/fail; instruction pass/fail; correction note. Below, record both minima and the combined decision.

D1: complete fictional samples for the async run

The following six text samples and readings are invented instructional inputs. They are not actual printed copies, students' observations or patient data. Use them to execute the simulated calculation branch, separately from writing the plan.

B1 sample: sheet 1 Q1 [blank], Q2 [blank]; sheet 2 Q3 [blank], Q4 [blank]; sheet 3 Q5 [blank].
Instructions: Write a question; Circle a priority; Bring the sheet; Leave a box blank if unsure.

B2 sample: sheet 1 Q1 [blank], Q2 [blank]; sheet 2 Q3 [blank], Q4 [blank]; sheet 3 Q5 [blank].

Instructions: Write a question; Circle a priority; Bring the sheet; Leave a box blank if unsure.

B3 sample: sheet 1 Q1 [blank], Q2 [blank]; sheet 2 Q3 [blank], Q4 [blank]; sheet 3 has no Q field.

Instructions: Write a question; Circle a priority; Bring the sheet. The omission is deliberately supplied.

A1 sample: one sheet, Q1 [blank], Q2 [blank]; instructions Write a question; Circle a priority. A2 sample:

one sheet, Q1 [blank], Q2 [blank]; instructions Write a question; Circle a priority.

A3 sample: one sheet, Q1 [blank], Q2 [blank]; instructions Write a question; Circle a priority. These are three named simulated copies, not independent participants.

Supplied readings, in copy order: B fields 5, 5, 4; B instruction types 4, 4, 3. A fields 2, 2, 2; A instruction types 2, 2, 2. Check each reading against the full sample text before entering it.

Async task: enter all six samples in the R2 log, compute both minima for each design, and determine each metric and combined result. State what should be repaired and why this does not establish clinical performance.

Calculation check after your attempt: did you take each design's minimum across all three copies for both metrics, compare each minimum with its R1 threshold, and pass the set only if both metrics pass? Recheck any entry against its full sample text.

Retain the original log, including any failed copy. For a failure, propose a repair to the reproduction process, then check a new version. Do not silently edit a sample or claim the rerun happened.

C1: warmups, three vague claims and support

Warmup: what evidence would support your specific prototype claim? Explain what a control is and why it helps. Rewrite "the device makes people more comfortable," or explain what evidence that claim would still require.

Prerequisite claim 1: "The sheet holds enough questions." Narrow rewrite: every B copy retains at least 5 distinct blank question fields. Unit: fields/copy. This counts structure, not question quality.

Prerequisite claim 2: "The directions are more complete." Narrow rewrite: every B copy contains all 4 S1 instruction types. Unit: types/copy. This does not show that readers understand them.

Prerequisite claim 3: "The device makes people more comfortable." Comfort is not measured by the supplied text. A safe replacement is the five-field structural claim, but state explicitly that this changes the claim and leaves comfort untested.

On-grade work includes the full two-metric plan, procedure, baseline and error fix. Use R2 as the pre-lab template. A missing prototype does not justify inventing observations.

Remediation: begin with one metric, one pass rule and A baseline. If you have an actual literature review, identify a relevant measurement unit and cite its source. Otherwise label R1 units as supplied instructional definitions.

Complete the second metric, procedure and error-reduction entry before calling the required plan complete. One-metric work is a partial scaffold. Missing authentic literature or required template remains pending.

If interrupted, find the first missing item in R2. Continue there. The simulated run requires D1 entries and computations; listing a plan for those calculations alone does not complete it.

X1: review, extension and final scope check

Peer review: can a partner reproduce your counts and decisions from the written definitions and sample records alone? Check both metrics, units, thresholds, repeat order, baseline and correction rule. Revise one unclear step.

If no partner is available, label a solo review. Hide earlier totals, recount D1 in reverse order and compare. Record your real correction or agreement; never invent a peer comment.

Extension: consider a blind review or sham comparison and name a remaining confounder. Masking A/B names may reduce expectation bias, but visible sheet and field differences limit masking.

A sham imitates a treatment or feature without the active element. It is not automatically useful for this text-count check and is not permission for participant testing. Explain whether it would help your actual question.

Number of sheets and instruction content both differ between A and B. They would complicate a broader claim about ease of use. A baseline and thresholds do not remove those differences or establish causal health benefit.

Self-check: "The prototype makes the patient healthier" lacks a defined outcome and acceptable evidence. A numerical change alone would not validate that broad claim. Rewrite as a bounded structural metric and name the health claim left untested.

Exit: name both metrics, their units and their pass thresholds. Distinguish plan complete, simulated calculation complete and actual validation pending. Keep the plan, simulation log and review notes in one record.

P1: actual course and weekly requirements

The date is Monday, April 5. Today, finish Problem 3.1.3 Design Innovations with two measurable metrics, in Class Records, and add the plan to the Problem 3 portfolio before leaving.

The broader project includes proposal, model/prototype or detailed schematic, and marketing pitch. Obtain the actual assigned portion, rubric and stopping point. This original plan and simulated log do not certify completion of that broader project.

Bring actual March 25 matrix/justification and chosen design when required. March 24's literature review, actual citations and official searches remain separate. Never claim these prior products are done because today's recovery is complete.

This week, April 5-9, includes prototype validation notes, testing and simulated validation planning. This class writes the plan today, Monday.

The broader unit also needs a problem statement, three source facts with peer-review status, matrix, two metrics, trial count and success rules. Actual prototype/model, materials, measuring tools and posted templates remain separate inputs when required.

No homework is due before this class. April 6 is catch-up. April 7 and April 8 have their own preparation items; do not relabel those as today's. The innovation-safety debate is an alternate requiring teacher direction.

The daily async route requires sample readings entered into a test-plan template and a pass/fail computation. D1/R2 supply a complete original route. The PhET landing page is optional extra practice; it does not replace the D1 calculation.

Use the teacher-confirmed Class Records assignment and portfolio route. Open the linked BI assignment in Turn in your work. Save locally if access is missing. Actual user or clinical validation needs a separately reviewed, authorized protocol; this lesson authorizes none.

M1: distinct worked plan for a paper document sleeve

Invented model: a paper sleeve is meant to fit a 300 by 220 mm document. The narrow claim is internal dimensions, not protection from damage or user satisfaction.

Metric L: minimum internal length across three sleeves, millimeters, pass at least 300 mm. Metric W: minimum internal width, millimeters, pass at least 220 mm. Both must pass. Set these rules before inspecting results.

Planned procedure: identify version and three sleeves; use the same ruler and millimeter scale; measure the inside from edge to edge twice per dimension. Keep the ruler straight and record disagreements before rechecking. An older sleeve design is the baseline.

Complete fictional readings: new sleeves N1 302 by 221 mm; N2 301 by 219 mm; N3 303 by 222 mm. Baseline O1, O2 and O3 are each 290 by 210 mm. No physical measurement occurred.

Calculation: new minimum length is 301 mm, so L passes; new minimum width is 219 mm, so W fails. The combined result fails. Baseline minima 290 and 210 mm fail both thresholds.

Error and fix: a slanted ruler changes the apparent length; align the ruler and repeat. Retain the failed width and propose a wider design. A proposed change is not a passed retest. Transfer this structure to B's field and instruction counts.

Try the decision

The D1 log records every B copy against both minimum-count rules. What should the writer conclude?

- Apply both minimum-count rules to every copy, retain the full log, propose a repair for any failure and keep real validation pending.
- Accept the simulated set if most samples meet both minimum-count rules.
- B is clinically effective because it contains more fields than A.

Explain your choice using the reading. Discuss it or say it to yourself. This practice does not add another submission.

Your 80-minute flow

0–5 minutes	Warm-up: what would count as proof that your prototype does what you claim?
5–20 minutes	State your prototype claim and choose two measurable metrics with units
20–40 minutes	Write the test procedure with pass/fail thresholds
40–55 minutes	Add a control or baseline and identify one error source
55–70 minutes	Peer review: can your partner run your test from your written plan alone? Apply the plan to D1 simulated readings and compute both results; use the solo route if needed.
70–80 minutes	Exit ticket: name your two metrics and their pass thresholds

Your response record

Use these labels on paper or in your document. The space below also works when this page is printed.

Define the claim

Answer all C1 warmups and repair the three vague claims. State actual design or explicit B recovery status.

A narrowed structural claim leaves comfort and health untested.

Use this space or the same labeled section in your existing document. Add paper when your table, graph or drawing needs more room.

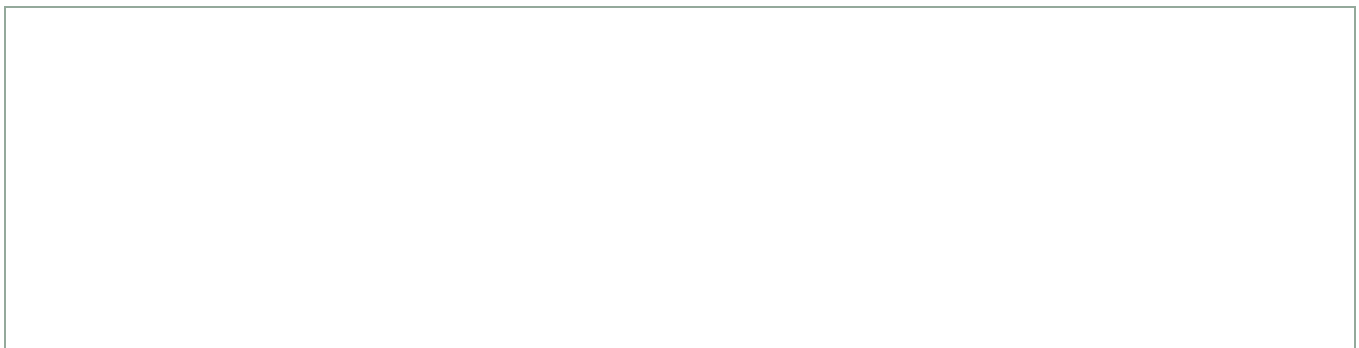


Choose two metrics

Write the structural claim, both metrics and units, and a numerical pass/fail rule for each. Explain why both matter to the stated requirement.

Use R1, or relevant actual-topic evidence if available. One metric is unfinished work.

Use this space or the same labeled section in your existing document. Add paper when your table, graph or drawing needs more room.

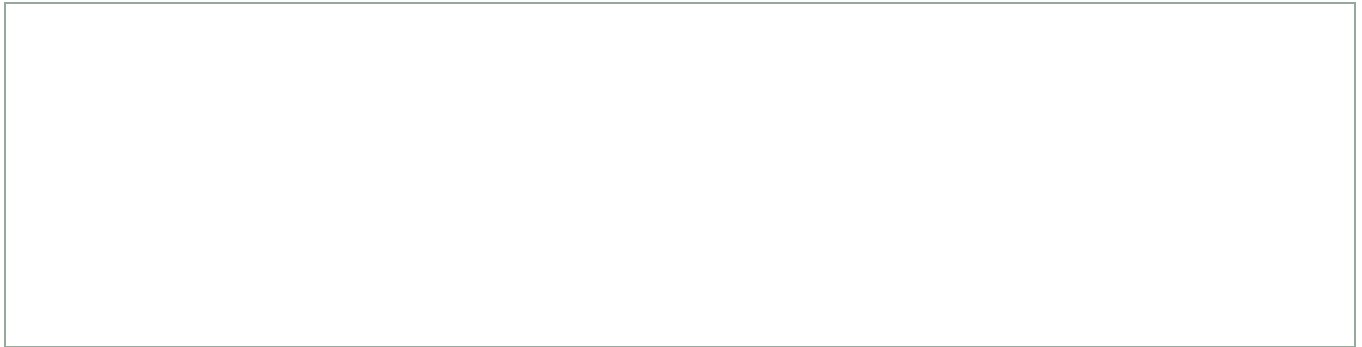


Write the test plan

Complete every R2 template entry. Describe three copies per design, ordered inspections, recording and the combined result rule.

Your plan must permit another reader to apply the same definitions without guessing.

Use this space or the same labeled section in your existing document. Add paper when your table, graph or drawing needs more room.

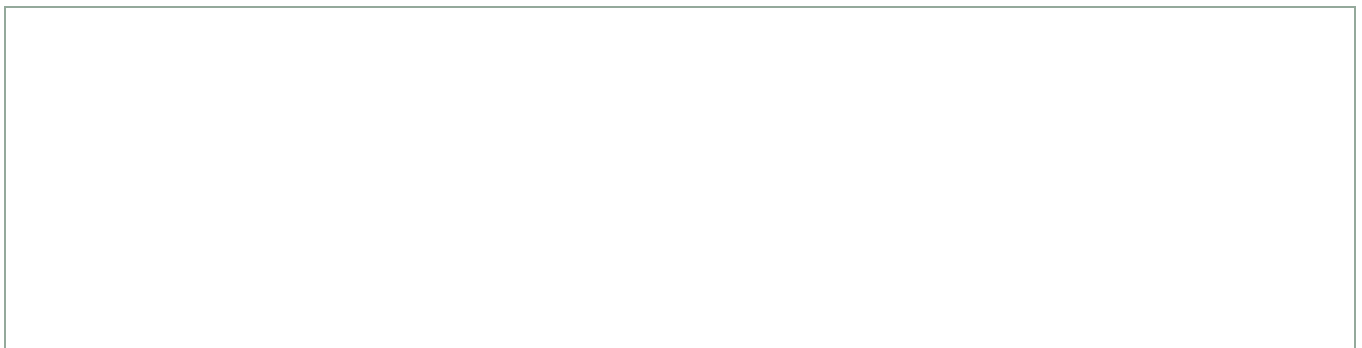


Specify the baseline and error fix

Name A or your documented actual baseline. State one likely error and a concrete reduction step; name what the comparison cannot establish.

A baseline is not proof of causation or medical benefit.

Use this space or the same labeled section in your existing document. Add paper when your table, graph or drawing needs more room.



Review and run the simulation

Use the partner or solo X1 check. Enter all six D1 samples in the R2 log; compute both minima and each pass/fail result. Revise one unclear step.

Label every supplied reading fictional. Keep failures and proposed repairs distinct from a completed rerun.

Use this space or the same labeled section in your existing document. Add paper when your table, graph or drawing needs more room.

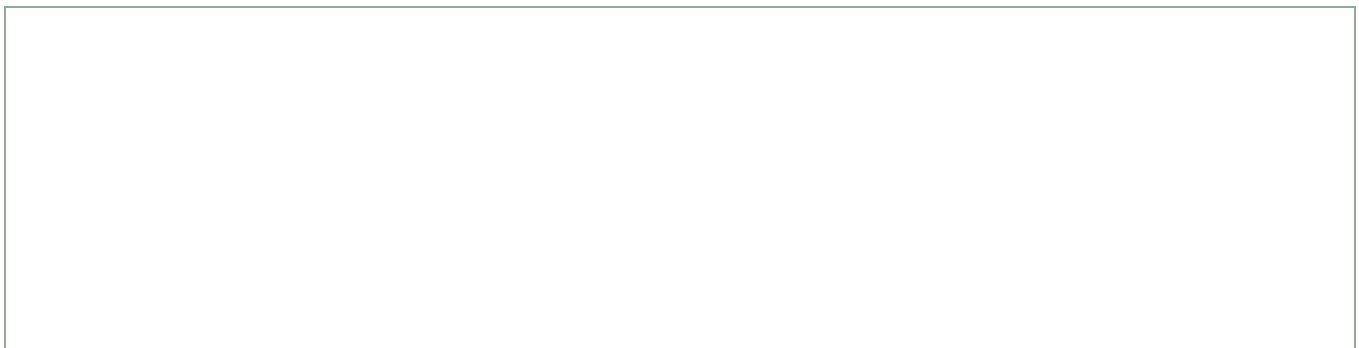


Finish and preserve status

Answer the X1 self-check and exit; add the optional blind/sham/confounder extension if ready. Submit the plan with its simulated log and review notes.

Use P1's confirmed route or local fallback. Actual official work and real validation remain pending until done.

Use this space or the same labeled section in your existing document. Add paper when your table, graph or drawing needs more room.



Check your product

- Your plan names two measurable metrics with pass/fail thresholds.
- You included a control and one error-reduction step.
- I preserve full sample logs and distinguish a plan, fictional computation and actual validation.

Keep the limit: Original paper-record requirements and simulated readings cannot establish usability, comfort, health benefit, clinical safety, causal effects or official project completion.

Turn in your work

Submit the validation plan through the teacher-confirmed Class Records assignment and add it to the assigned Problem 3 portfolio. Keep the simulated log and review notes with the plan. Confirm actual assignment access and stopping point.

Save the same record locally. List the missing assignment, authentic design, sources or protocol without claiming completion.

[Open this assignment in Schoology.](#)

Carry forward

Set the rule before the result, keep failed samples, and limit the claim to what the check measures.

Retain the plan and simulated log; obtain actual project directions before any further testing. What result would count as evidence that the proposed design meets its criterion?

Sources and resources

[Open today's page](#)

NASA-VV · Systems Engineering Handbook, 5.0 Product Realization

NASA

<https://www.nasa.gov/reference/5-0-product-realization/>

NIST-OBJ · 5.3.1 What are the objectives?

NIST/SEMATECH e-Handbook of Statistical Methods

<https://www.itl.nist.gov/div898/handbook/pri/section3/pri31.htm>

FDA-HF · Human Factors and Medical Devices

U.S. Food and Drug Administration

<https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/human-factors-and-medical-devices>