FIELD GUIDE · INNOVATION PRACTICE

Developing assumptions from a proposed solution.

A working method for turning a proposed solution into clear, testable assumptions across the five IXD dimensions, and rating evidence honestly before scale.

IXD INNOVATION GUIDE · EDITION 01 · 2026

PURPOSE

The goal is not to prove the solution is right.

The goal is to identify what must be true for the solution to work, then build confidence through evidence. Every solution rests on assumptions across five IXD dimensions.

01
Desirability

Do users, customers, or team members want or need this?

02
Operational feasibility

Can the business actually operate this effectively?

03
Technical feasibility

Can we build, integrate, run, and support it?

04
Viability

Does it make commercial, strategic, or value sense?

05
Adoption readiness

Will people adopt, use, and sustain the change over time?

Step 1–3	Step 4–6	Step 7–9
Frame the solution and map assumptions	**Write assumptions, rate evidence honestly**	**Prioritise, test, decide what evidence means**

STEP 1

Express the solution and articulate the hypothesis.

Make the proposed solution clear enough for people to understand, challenge, and test. Then express the hypothesis it depends on.

USEFUL WAYS TO EXPRESS A SOLUTION

Service blueprints
People, process, systems, handoffs, backstage activity.

Concept expressions
The idea, value proposition, and intended experience.

Storyboards
How the solution fits a real user or team journey.

Sketches
Quick visuals that make an idea discussable.

It does not need to be polished. It needs to be specific enough that people can ask: **what would need to be true for this to work?**

SOLUTION STATEMENT

We believe that **[solution]** will help **[users / teams]** to **[achieve outcome]** by **[how it works]**.

We believe that an automated exception triage dashboard will help store teams resolve online order issues faster by prioritising exceptions based on age, value, and fulfilment risk.

HYPOTHESIS STATEMENT

If we **[introduce solution]**, then **[users / teams]** will **[change behaviour]**, resulting in **[measurable benefit]**.

If we introduce an automated exception triage dashboard, then store teams will resolve the highest-risk online order issues earlier in the day, resulting in fewer aged exceptions and customer escalations.

STEP 2

Identify what must be true.

For this solution and hypothesis to hold, what are we assuming is true? Write assumptions as statements, not questions.

INSTEAD OF (QUESTION)

Will store teams use it?

Can IT build it?

›

WRITE (STATEMENT)

Store teams will use the dashboard daily as part of their exception resolution workflow.

The required exception data is available, reliable, and can be integrated into a dashboard.

A USEFUL ASSUMPTION IS

Specific enough to test, and important enough that being wrong would change the solution, hypothesis, or decision.

STEP 3 · DIMENSION 01

Desirability.

Focus. Do people need this, want this, or value this?

What to ask

Who is the user or customer?

What problem are they experiencing?

How painful or frequent is the problem?

Would the solution make their work or experience meaningfully better?

What behaviour are we assuming they will change?

Evidence to look for

- User interviews
- Store feedback
- Customer complaints
- Support tickets
- Time-in-task observation
- Workaround analysis
- Survey responses
- Prototype usage

Example assumptions

Store teams experience exception resolution as a high-friction task.

Store teams want a prioritised view rather than a full list of exceptions.

Team members trust system-generated prioritisation.

Customers will benefit from faster exception resolution.

The proposed solution addresses a problem important enough to act on.

STEP 3 · DIMENSION 02

Operational feasibility.

Focus. Can the business actually operate this solution?

What to ask

Who will own the process?

What changes to roles, routines, or handoffs are required?

Can this fit into current operating rhythms?

What happens when the solution fails or creates an exception?

Are there policy, compliance, training, or support implications?

Evidence to look for

- Process walkthroughs
- Store trial feedback
- Role mapping
- Capacity checks
- Exception scenarios
- SOP / policy review
- Pilot performance data
- Observed real-store use

Example assumptions

Store teams have capacity to review and act on prioritised exceptions daily.

The new process fits into existing store operating rhythms.

Clear ownership exists for unresolved or escalated exceptions.

Support teams can manage questions or failures without significant extra workload.

The process can be standardised across store formats or volumes.

STEP 3 · DIMENSION 03

Technical feasibility.

Focus. Can we build, integrate, run, and maintain it?

What to ask

What systems, data, integrations, or platforms are required?

Is the data available, accurate, timely, and usable?

Are there dependencies on other teams or roadmaps?

What are the security, privacy, architecture, or support considerations?

Can the solution perform reliably at the required scale?

Example assumptions

Required data fields are available from existing systems.

Data quality is sufficient to support prioritisation decisions.

Integration with current systems is technically achievable.

The solution can meet performance and reliability expectations.

Ongoing support and maintenance can be handled by existing teams.

Evidence to look for

- Technical spike
- Data profiling
- Architecture review
- Integration assessment
- Security / privacy review
- Vendor assessment
- Working prototype / MVP
- Production-like load test

STEP 3 · DIMENSION 04

Viability.

Focus. Is the solution worth doing?

What to ask

What value does this create?

Is the value measurable?

What costs, trade-offs, or opportunity costs are involved?

Does it align with business priorities?

What would make this solution not worth pursuing?

Has Finance co-designed the viability test where scale decisions are involved?

Evidence to look for

- Business case
- Cost estimate
- Benefit sizing
- Baseline metrics
- Value driver analysis
- Prioritisation comparison
- Finance co-designed test
- Measured business impact

Example assumptions

Faster exception resolution will reduce cancelled orders or customer dissatisfaction.

The solution will save enough time or cost to justify investment.

The benefit is large enough compared with other opportunities.

The solution supports current business priorities.

Ongoing operating costs will be acceptable.

The value case holds beyond initial novelty or pilot enthusiasm.

STEP 3 · DIMENSION 05

Adoption readiness.

Focus. Will people adopt and sustain the change?

What to ask

Who needs to change behaviour?

What training, communication, or reinforcement is needed?

Are users motivated and able to adopt the solution?

What resistance might appear?

How will adoption be measured?

Did people actually use it in realistic operating conditions, or did they only say they would?

Evidence to look for

- Change impact assessment
- Training feedback
- Pilot adoption metrics
- Usage analytics
- Leader feedback
- Comms testing
- Behaviour observation
- Sustained usage across stores

Example assumptions

Store leaders will reinforce use of the new workflow.

Team members can learn the solution with minimal training.

Users will stop using existing workarounds.

Adoption can be monitored through usage or process metrics.

The change can be embedded into standard operating routines.

Usage will continue after the novelty effect wears off.

STEP 4

Write assumptions in a testable format.

A good assumption is specific, testable, linked to one of the five dimensions, connected to the hypothesis, and important enough that being wrong would change the solution or decision.

We assume that **[user / system / team / process]** will **[behaviour / outcome]** because **[reason]**.

SUBJECT
The user, system, team, or process whose behaviour is at stake.

BEHAVIOUR
What they will do, observe, or produce that matters for the solution.

REASON
Why we think this will happen. The argument we are testing.

TEST
If the reason is false, the behaviour likely changes. That is what we test for.

EXAMPLE

We assume **store teams** will **check the dashboard daily** because it gives them a faster way to identify the most urgent exceptions.

EXAMPLE

We assume **exception data** is **accurate enough for prioritisation** because the source systems already capture exception age, order value, and fulfilment status.

EXAMPLE

We assume **store leaders** will **support the process** because it helps reduce customer escalations and manual follow-up.

STEP 5

Rate confidence using the evidence scale.

Each assumption starts as **Assumed**. The rating describes the strength of evidence, not the team's optimism. If we only have what people said, the highest rating is Indicative.

L1 **Assumed**	**Belief only.** Hypothesis, expert opinion, untested logic. Every new assumption starts here. Honest, not shameful.
L2 **Indicative**	**Signals suggest it may be true.** Interviews, surveys, desk analysis, analogous examples. *What people say they would do is always Indicative at best.*
L3 **Demonstrated**	**Observed behaviour in realistic operating conditions in one real store.** First level that counts as behavioural evidence. Plausible becomes "we saw it."
L4 **Confirmed at scale**	**Decision-grade.** Representative multi-store trial, sustained usage past novelty, Data and Finance co-designed evidence. Suitable for scale or investment decisions.

OFF-RAMP AVAILABLE AT ANY LEVEL

↓ **Disproven**	**A fair test contradicts the belief.** Report with pride. It means a downstream mistake was avoided.

STEP 6

Apply the trigger question.

Have we observed actual behaviour in realistic operating conditions?

IF NO		OR	IF YES	
Assumed	no evidence yet		**Demonstrated**	one real store
Indicative	signals only · what people said		**Confirmed at scale**	representative trial

Disproven · fair test contradicts the belief

STEP 7

Prioritise using a 2×2 matrix.

Plot each assumption by **risk if wrong** against **current evidence**. The riskiest, least-evidenced quadrant earns the next test.

	LOWER EVIDENCE **Assumed**	HIGHER EVIDENCE **Demonstrated**
High risk IF WRONG	**Critical assumptions** TEST FIRST Could materially change the solution, hypothesis, investment, or delivery.	**Monitor / strengthen for scale** TEST BEFORE SCALING Evidence exists, but may need broader testing before commitment.
Low risk IF WRONG	**Park or test lightly** ONLY IF CHEAP / NEAR-TERM Test only if quick, cheap, or needed for an upcoming decision.	**Use as supporting evidence** REUSE, DON'T RETEST Do not spend more effort unless conditions change.

INSIDE THE TOP-LEFT QUADRANT

High risk if wrong. Currently only Assumed.

Critical assumptions have the greatest potential to change the direction of the work. Test them before lower-risk or better-evidenced ones.

Examples

Store teams will use the dashboard daily without adding workload pressure.

The data required to prioritise exceptions is available and accurate enough.

Faster exception resolution will reduce customer escalations.

Store leaders will reinforce the new workflow after the trial period.

The solution can operate within existing store routines.

When several are critical, ask in this order

01 Would this disprove the core hypothesis?

02 Would this change the proposed solution?

03 Would this change whether we should invest?

04 Would this change the operating model or ownership?

05 Would this change whether people will adopt it?

RULE OF THUMB

The first test targets the assumption with the highest potential to change what the team does next.

STEP 8

Convert critical assumptions into experiments.

For each priority assumption, define how it will be tested, what evidence is needed, and what would count as a success signal.

ASSUMPTION	DIMENSION	EVIDENCE	RISK	TEST METHOD	SUCCESS SIGNAL
Store teams will use the dashboard daily	ADOPTION	ASSUMED	HIGH	Two-week pilot with one store	Used on 80% of trading days
Data is reliable enough for prioritisation	TECHNICAL	ASSUMED	HIGH	Compare dashboard data against source records	95%+ match on key fields
Faster resolution reduces cancellations	VIABILITY	ASSUMED	HIGH	Compare pilot store to baseline	15% reduction in aged exceptions
Daily usage continues after launch excitement	ADOPTION	DEMONSTRATED	HIGH	Representative multi-store trial	Usage stable after novelty effect
Value case holds at scale	VIABILITY	DEMONSTRATED	HIGH	Data and Finance co-designed scale trial	Benefits exceed agreed threshold

DESIGN THE EXPERIMENT TO MOVE CONFIDENCE

ASSUMED → INDICATIVE · Learn from people, data, or prior examples

INDICATIVE → DEMONSTRATED · Observe behaviour in realistic operating conditions

DEMONSTRATED → CONFIRMED AT SCALE ·

Decision-grade evidence across representative stores

STEP 9

Synthesize the evidence and decide what it means.

After testing, do not just update the confidence rating. Synthesize what the evidence means for the assumption, hypothesis, and solution.

01 **Does this support or challenge the assumption?**

Make it more believable, less believable, or disprove it?

02 **Does this support or challenge the current solution?**

On the right path, needs adjustment, or should be reconsidered?

03 **Does this change our mind about the hypothesis?**

Is the original hypothesis still plausible, or does it need reframing?

04 **Does this introduce new data to consider?**

New constraints, behaviours, costs, risks, or dependencies?

05 **Does this create new critical assumptions?**

Did the evidence reveal a new high-risk belief still only Assumed?

06 **What decision should we make now?**

Proceed, pivot, stop, learn more, or scale.

EVIDENCE SYNTHESIS TEMPLATE

What did we test · what evidence did we gather · was it behavioural · which assumption did it support or challenge · previous and updated confidence · effect on the solution and hypothesis · new constraints · new critical assumptions · decision.

HOW TO INTERPRET EVIDENCE

Four modes, four next moves.

Supports the assumption

- Is the evidence behavioural or only stated intent?
- From one store or representative stores?
- Is the effect strong enough to matter?
- What becomes the next priority?

THEN Confidence moves up. Decide if scale evidence is still needed before investment.

Partially supports

- Which part was supported, which was not?
- Was the issue solution, context, process, data, or adoption?
- Does the solution need to change?
- Is there a new critical assumption?

THEN Adjust the solution or surface a new critical assumption to test next.

Challenges the assumption

- Was the test fair?
- Did users behave differently than expected?
- Did operations, systems, or incentives interfere?
- Does this challenge the solution or the hypothesis?

THEN Pivot, stop, or run a different test that isolates the real cause.

Disproves the assumption

- What belief did we avoid carrying forward?
- What downstream decision did this protect us from?
- Does this disprove the solution, or only one version?
- What should stop now, what should be redesigned?

THEN Mark Disproven, stop the affected work, redesign or revisit the value case.

DECISION OPTIONS

Five ways to act on the evidence.

Proceed
Evidence supports the assumption strongly enough for the next decision.

Pivot
Problem or need stands; the current solution must change.

Stop
Evidence contradicts the assumption, solution, or value case.

Learn more
Evidence is inconclusive, too limited, or not yet behavioural.

Scale
Confirmed at scale, decision-grade.

WORKED EXAMPLE

Example synthesis across four assumptions.

ASSUMPTION	PREVIOUS	NEW EVIDENCE	UPDATED	DECISION
Store teams will use the dashboard daily	ASSUMED	Used on 9 of 10 trading days in one real store	DEMONSTRATED	**Proceed** to broader trial
Data is reliable enough for prioritisation	ASSUMED	87% match on key fields	INDICATIVE	**Pivot** data approach
Faster resolution reduces cancellations	ASSUMED	No measurable change in pilot	DISPROVEN	**Stop** · revisit value case
Teams sustain daily usage across store types	DEMONSTRATED	Representative trial shows stable usage after novelty	CONFIRMED	**Scale**

COMMON MISTAKES TO AVOID

Eight ways teams get this wrong.

Mistake 01

Writing risks instead of assumptions

Risk: Stores may not use the dashboard. **Assumption:** Store teams will use the dashboard daily to prioritise exception resolution.

Mistake 02

Writing vague assumptions

Vague: Data will be good. **Better:** Exception data is accurate enough to prioritise with fewer than 5% material errors.

Mistake 03

Treating interviews as behavioural evidence

What people say they would do is always Indicative at best. Demonstrated requires observed behaviour in realistic operating conditions.

Mistake 04

Prioritising by ease instead of risk

Incorrect: Easiest to test, so test first. **Better:** High risk if wrong and only Assumed, so it is a critical assumption.

Mistake 05

Testing too many critical assumptions at once

When several are critical, focus the next test on the one most likely to change the solution, investment, operating model, or adoption path.

Mistake 06

Only focusing on technical assumptions

A solution can be technically possible and still fail because users do not want it, operations cannot absorb it, value is too small, or adoption is too hard.

Mistake 07

Treating assumptions as facts

Until tested, assumptions are beliefs. Use the confidence rating to show evidence strength.

Mistake 08

Avoiding Disproven

Disproven is not failure. It means the team learned before making a bigger, more expensive mistake.

SIMPLE WORKING TEMPLATE

One row per assumption.

Use this template in any working doc to track assumptions through the process. Update the confidence rating honestly as evidence is gathered.

DIMENSION	ASSUMPTION	EVIDENCE TODAY	EVIDENCE LEVEL	RISK	CRITICAL?	TEST / NEXT STEP
Desirability	*e.g. Store teams want a prioritised view.*	*12 store interviews.*	INDICATIVE	HIGH	YES	*Observe usage in one store.*
Operational	…	…	ASSUMED	LOW	NO	…
Technical	…	…	ASSUMED	HIGH	YES	…
Viability	…	…	ASSUMED	HIGH	YES	…
Adoption	…	…	DEMO	HIGH	NO	…

One-line rule. A good assumption is specific, testable, linked to one of the five dimensions, rated honestly based on evidence, and prioritised by how much it would change the solution if wrong.