CarryWell

One student idea, run through Naventros from Stage 0 to Stage 5

Two students on a Dubai campus start with an idea for a reusable lunch container. Twenty-two weeks later they have changed the problem, dropped a feature that failed its test, and are ready to make their first sale. This is how the method got them there.

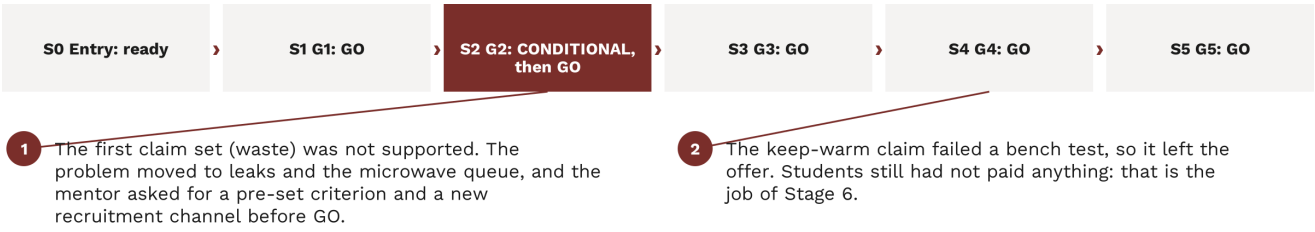
Example startup journey
Fictional case: people, campus, figures and research results are invented for teaching
Naventros, October 2026

About this example

> **Everything in this example is fictional.** Layla and Omar, their campus, the people they interview and every figure and research result, including what online sources "showed", are invented to show how the method works. Nothing here is market data about real products, platforms or universities.

The founders. Layla is a second-year business student who has run a stall at a campus fair. Omar is a third-year engineering student who commutes 50 minutes each way and brings lunch from home most days. They meet in the first Naventros session and form a team.

How to read it. Each stage shows what the founders did, what they found, how Your AI Partner helped and where it was overruled, and what the human reviewer decided. The weeks follow the 30-week Naventros calendar.



CarryWell's review path through the core programme.

Stage and weeks	What happened	Review
0 Explore an opportunity (weeks 1-3)	A product idea is reframed into a problem worth checking	Entry check: ready
1 Founder and opportunity (weeks 3-5)	Claims written so that they could fail; their own AI assistant briefed	G1: GO
2 Customers and value (weeks 6-9)	Three investigations; the first claim set fails; the problem moves	G2: CONDITIONAL, then GO
3 Solution and value (weeks 10-12)	Twenty-four ideas, ranked against criteria written first	G3: GO
4 Business model (weeks 13-15)	The supplier's minimum order meets the founders' affordable loss	G4: GO
5 Prototype and tests (weeks 16-22)	Three tests with thresholds set beforehand; one claim fails	G5: GO on method; demand evidence weak

Stage 0: Explore an opportunity

Weeks 1 to 3. The question: what is worth our next few weeks, and why us?

Where they started. Layla arrived with an idea: "a stylish reusable lunch container for students, to cut single-use waste". In the first session the facilitator asked every student to write their idea down and then to keep a seven-day friction journal instead of working on it.

What they noticed. Their journals and one observation morning in the student lounge produced twenty frictions. Clustered, the strongest group was about eating on campus: soup leaking into Omar's laptop bag on the metro; a queue of fourteen people at the lounge's two microwaves; classmates skipping lunch to save money; takeaway tubs overflowing the bins.

Reframe. Using the 5R Reframe worksheet, they found that Layla's sentence was a solution in disguise: it named a product before it named a problem, and it assumed waste was what students cared about.

BEFORE

"Students need a reusable lunch container to cut single-use waste."

- − A solution in disguise: it names the product before the problem.
- − Assumes waste is what students act on.
- − Nothing in it can be checked with a person this week.

AFTER

"Students who bring lunch from home three or more days a week lose meals, money and time to leaks in their bags and to the microwave queue."

- + A named group and a situation that can be observed.
- + Costs that can be counted: a ruined meal, a bought lunch, minutes queued.
- + Came from interviews, not from the founders' first idea.

The right-hand statement is where they ended up in week 9, not week 3. In week 3 they committed to a broader problem: *students who eat lunch on campus struggle to eat well, cheaply and without waste between classes*, with three first-hand observations behind it.

Your AI Partner. Asked to critique the problem statement, it pointed out that it bundled three problems (cost, quality, waste). The founders kept the bundle for now and wrote in their decision log that Stage 2 would separate them. **Entry check: ready.**

Stage 1: Founder and opportunity

Weeks 3 to 5. The question: what are we trying to do, with what means, and what could be wrong?

Means and affordable loss. Omar has access to the engineering lab and lives the commuter's lunch problem; Layla has contacts in three student societies and stall experience. They agreed what they could afford to lose: six hours a week each until the end of the year and 500 CU each, 1,000 CU in all.

Claims that could fail. They wrote their beliefs as claims with criteria set before any research.

Claim	Their first version (first claim set)	Criterion, written in week 4
C1 The need exists	Students who eat lunch on campus are bothered by single-use packaging	Supported if a clear majority of valid students describe something specific they did in the last 30 days to avoid single-use packaging
C2 It matters enough	They would act to avoid it	As C1, with a cost in money or time
C3 They already spend	They spend money or effort on alternatives today	Supported if several report a specific spend on an alternative

Your AI Partner. They set up the venture briefing page and gave it to the assistant. It proposed "health-conscious young professionals" as a second segment. They **rejected** it: no observation supported it, and it would have split a two-person team. Their AI use note recorded the suggestion, the rejection and the reason.

G1, Opportunity profiled: GO. The reviewer checked that each claim could be shown false and that the criteria were written before the data.

Stage 2: Customers and value

Weeks 6 to 9. The question: does a defined group really have this need, and what do they do and spend about it today?

The research plan: three investigations first

Your AI Partner proposed a research plan of eight items, tied to the claims and to the decision the founders had to make. The founders chose the first three, for two visible reasons: the claim their decision depended on (C1, C2), and the claims with the least evidence in their register. The other five stayed on the list.

Investigation	Method	What it can show	What it cannot show
1. How people search and what they complain about	Search-interest tool (public interface, five years, national level, downloaded file); the top twenty listings in the lunch-box category on two national online marketplaces, read by hand with dated screenshots; sixty recent one- and two-star reviews, copied as passages	Language, seasonality, price bands, recurring complaints	Buying demand, sales volumes, or anything about campus students in particular
2. What students actually do at lunch	Counting at the main lounge across three lunch periods; timing the microwave queue	Behaviour of a defined group, observed	Why they do it; anything about other campuses
3. What students say about the last time	Interviews using a guide built in lesson U2.3, consent recorded with five permissions	Specific past events, workarounds, spend	Whether they would buy a new product

What they found

Investigation	Finding (fictional)	Recorded limits
Search interest	Interest in lunch-box terms rose around the start of each academic year and in January; "leak proof" appeared among related searches	Relative interest only, national, not students, not buying
Marketplace listings	Basic plastic boxes at 15-35 CU; insulated and steel boxes at 60-140 CU; bestsellers were large multi-compartment boxes	A ranking is not a sales figure; buyers are not necessarily students
Low-star reviews (60)	Leaks 23; broken lids or latches 14; too bulky 9; smell or staining 7; other 7	Self-selected reviewers; complaints about products, not about campus life
Lounge observation (412 lunches)	96 home-made lunches in a container (23%), 251 canteen or takeaway (61%), 65 none visible (16%); 12 of the 96 were reused takeaway tubs	One lounge, three days, one week of term

Investigation	Finding (fictional)	Recorded limits
Microwave queue (20 timed waits)	Median wait 9 minutes; longest 22; up to 14 people queuing at two microwaves	Lunch peak only
Interviews, round 1 (14)	10 students who bring lunch from home three or more days a week; 2 who rarely do (comparators); 2 canteen staff (context)	All students recruited through one channel, the engineering society

Consent. All fourteen agreed to research notes. Two declined recording, so notes only. One declined external AI analysis: her notes were coded by hand and never entered the AI Partner's context. Three agreed to anonymous public quotation; none to named quotation.

Ledger excerpts (fictional).

#	Who	Statement or observation	Claim	Level	Source
2	Commuter, home lunch 5 days	"Tuesday last week my lentil soup opened in my bag on the metro. My notes were ruined and I bought lunch here instead, 22 CU."	C1, C2 (revised claim set)	E3	Reported
2	same	"Waste matters to me, definitely."	C1 (first claim set)	E1	Reported
5	Home lunch 4 days	"I bought a second box after the first one leaked, about 45 CU."	C3	E5	Reported
7	Home lunch 3 days	"If the queue is long I just eat it cold. That was twice this week."	C1, C2 (revised claim set)	E3	Reported
9	Home lunch 5 days	Showed a photo, dated eight days earlier, of a bag stained by leaked curry	C1, C2 (revised claim set)	E4	Artifact
12	Rarely brings lunch	"It's fine, I just buy something."	Comparator	E3	Reported

The first claim set fails

Against the criterion written in week 4, only 2 of the 12 students described anything they had done in the last 30 days to avoid single-use packaging. Eleven said waste "matters" (opinions, E1). **C1 and C2 for first claim set: not supported, moderate confidence.**

A different pattern was in the same notes. Of the ten students who bring lunch from home three or more days a week, seven described a specific failure in the last fourteen days: four leaks in a bag (one ruined notes, one stained a laptop sleeve, two lost the meal) and three gave up on the microwave queue (one ate cold, two bought canteen lunch at about 20 CU). Six of the ten had spent money on a workaround: a replacement box after a leak (30-60 CU) or a canteen lunch on queue days.

Revised claim set, written during the analysis. Population: students who bring lunch from home three or more days a week. Criterion: a clear majority describe a specific carry, keep or heat failure in the last 14 days that cost them a meal, money or more than ten minutes. Because it was written after seeing the data, it is marked **after the data**. Seven of ten met it.

The review

Step	Mentor's ruling
Integrity	Clear. Consent recorded for all fourteen; the interview without external AI permission was handled by hand
Stage test	First claim set not supported (moderate confidence). Revised claim set supported with **low** confidence: criterion written after the data, one recruitment channel, ten items, one campus
Status	**CONDITIONAL** for 30 days. Action: pre-register the revised criterion; recruit through at least one new channel; include at least one observed or artifact item
Learning result	Pass for both founders: method sound; the failed claim set recorded honestly

Continuation. In weeks 9 and 10 they recruited seven more students through a commuter-students society and by approaching people at the library café. Five of seven met the pre-registered criterion; one reported no failures (an insulated bag works for her); one had a leak with no cost. Three students showed dated photos of leaked bags. The contradiction with the online data was recorded rather than averaged away: bestsellers are large multi-compartment boxes, but the students they watched carry small single containers or reused takeaway tubs that fit a laptop bag.

Re-review, week 11: GO. C1 and C2 supported with moderate confidence for students who bring lunch from home three or more days a week (limits: one campus, mostly engineering and business students). C3 supported with low to moderate confidence.

How big is this, on one campus?



Every number is labelled. The biggest lever is the share who bring lunch three or more days a week, so that is the next research question.

The founders' conclusion: a small but real beachhead of a few hundred reachable students on one campus, worth testing; anything larger depends on later evidence from other campuses.

Where Your AI Partner helped, and where it was overruled

What the AI did	What the founders decided
Critiqued the interview guide: flagged three leading questions	Accepted all three rewrites; rejected its suggested 1-10 satisfaction rating (a rating is an opinion)
Proposed categories for coding the sixty reviews	Checked 20 of the 60 against the categories by hand and changed 4 codes
Proposed tags (claim, level, source) for interview notes it was permitted to see	Verified every tag used in a claim decision; changed 3
Offered a national market size for lunch boxes with no source	Not recorded. Labelled an unverified lead; they had no way to check it and it said nothing about their campus

Stage 3: Solution and value

Weeks 10 to 12. The question: which solution creates value that is clearly different, and how will we test it?

Range before judgement. In the Reimagine session they generated twenty-four concepts without judging any of them: a slim leak-tested container with an insulated sleeve; a deposit-tub scheme run with the canteen; a live display of the microwave queue; an insulated pouch for the box students already own; heated lunch lockers; and nineteen more.

Criteria first, then scores. Before scoring, they wrote five criteria and weights: fits the evidenced pain (weight 2); testable by the two of them within twelve weeks; test cost under 300 CU; distinct from what students already have; responsible (food safety and waste).

Concept	Weighted score (out of 30)	Note
Slim leak-tested container and sleeve	24	Chosen to test first
Canteen deposit-tub scheme	21	Rises to first if "distinct" is weighted double; kept as the partnership option
Pouch for the student's own box	18	Cheap, but does not fix leaks
Microwave queue display	16	Addresses heating, not carrying; parked
Heated lunch lockers	11	Not testable within twelve weeks

Value proposition. *For students who bring lunch from home, CarryWell is a slim container and sleeve that is tested not to leak in a laptop bag and keeps food warm until lunch, so you stop losing meals and queuing for the microwave.* The keep-warm promise was written as a claim to be tested (C7), not as a fact.

Testable hypotheses. C5: at least 15% of students in the segment who stop at a lunch table will register intent at 49 CU, with their contact details, within one week. C4 and C5 in use: at least 8 of 12 loaner users will use it three or more days a week for two weeks.

Positioning reactions. Eight students reacted to three positioning lines. The eco line fell flat; "never leaks in your bag" drew the most interest. Recorded as reactions (opinions), not as evidence of demand.

G3, Solution hypothesis defined: GO.

Stage 4: Business model

Weeks 13 to 15. The question: can this become a viable, responsible business on paper, and which assumption could sink it?

The model. Sales at lunch tables and to student societies in bulk. Because the founders hold no trading licence, sales must run through a licensed route such as the university's enterprise unit; this is a question for the university and its local adviser (see the country addendum), and the model assumes a 10% fee for it.

Line	Figure (fictional)	Label
Price	49 CU	Hypothesis, between basic boxes (15-35 CU) and premium ones (60-140 CU)
Unit cost, container and sleeve, landed (Supplier A, minimum order 100)	22 CU	Supplier quote
Payment fee and packaging	3 CU	Assumption
Enterprise-unit fee (10% of price)	4.90 CU	Assumption, to confirm with the university
Contribution per unit after all variable costs	about 19 CU	Calculated

The cash lesson. The smallest order from Supplier A is 100 units: 2,200 CU paid before a single sale, more than double the 1,000 CU they could afford to lose. Each sale returns about 41 CU of cash (49 CU less fees and packaging), so they would need to sell 54 of the 100 just to get their money back. In the stress test, if only 30 sold, they would be about 970 CU down: almost all of their affordable loss on one batch. Supplier B was cheaper per unit (15 CU) but required 500 units (7,500 CU) and was ruled out.

The riskiest assumptions. Importance and evidence were each scored 1 to 5 in the register.

Assumption	Importance	Evidence	Next step
C6 Students will pay 49 CU	5	1	Cannot be tested with money until a licensed route is confirmed; test intent first
C7 The container does not leak in a bag	5	2 (supplier's claim)	Bench test
C7 It keeps food warm until lunch	4	1	Bench test
C4 Students will switch from what they use now	4	2	Loaner pilot
A licensed selling route exists	5	1	Question to the university

Assumption	Importance	Evidence	Next step
The product is safe for food contact	5	1	Supplier certificate, checked by an adviser

Good business design. They reviewed harms: a new plastic product versus the waste it replaces (they chose a design the supplier states is recyclable, to be verified); food safety; and student budgets (no pressure selling, a clear price).

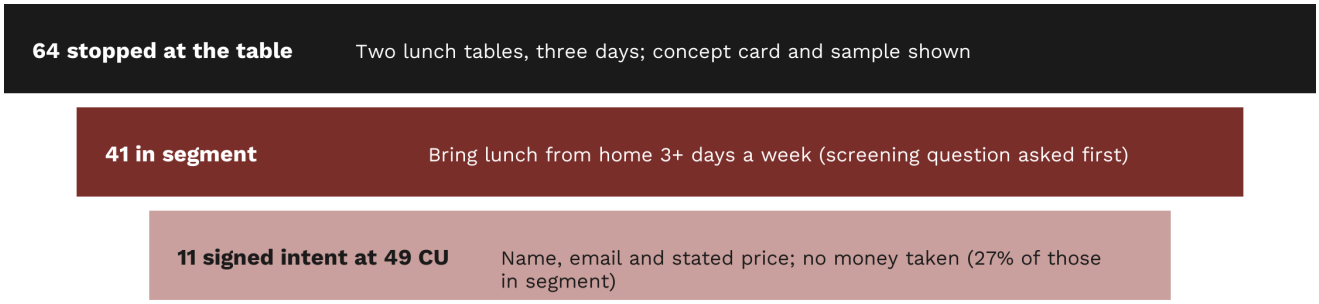
G4, Model feasible on paper: GO. The reviewer noted that the model was coherent, every key number was labelled, and the riskiest claims had test plans.

Stage 5: Prototype and tests

Weeks 16 to 22. The question: will real people act?

Each test had a Test Card approved by the facilitator before it ran, with the threshold written in advance.

Test	Claim	Threshold, set before	Result	Reading
T1 Bench test, 5 sample units (110 CU)	C7 leaks	No leaks in 5 units, soup, upside down in a bag for 2 hours	0 of 5 leaked	Supported
T1 Bench test, same units	C7 keeps warm	At least 4 of 5 above 50 °C after 4 hours	2 of 5	**Not supported**
T2 Intent table, two tables, three days	C5	At least 15% of in-segment students who stop sign intent at 49 CU; against if under 5%	11 of 41 (27%)	Supported, weak: stated intent with contact details
T3 Loaner pilot, 12 students, 2 weeks	C4, C5 in use	At least 8 of 12 use it 3 or more days a week	9 of 12; one leak (lid not fully closed)	Supported



The intent table. No money was taken: the founders had no licensed route to sell yet.

What the failed test changed. The keep-warm promise left the offer. The value proposition became: *a slim container and sleeve, tested not to leak in a laptop bag*. The heating problem, which the evidence showed was real, was recorded as unsolved, and the microwave-queue concept went back on the list. Six students at the tables asked for a cheaper version at about 35 CU; that was recorded for the pricing decision in Stage 6.

Your AI Partner. It drafted the concept card for the tables; the founders removed its "keeps food hot for hours" line, which their own test contradicted. It summarised the loaner check-in messages that participants had permitted for AI analysis; the founders checked every summary line against the messages before using it.

Decision memo: persevere, with a narrower offer. Next: a first paid sale of 40 units through the university's licensed route, to test C6 with money.

G5 GO on method. The stage test asks whether the experiments were run properly (thresholds set beforehand, raw data, a decision memo, a record of AI use), not whether they succeeded. The results are recorded separately: C5 supported (weak); C7 leak-proof supported, keep-warm not supported; C6 not yet tested.

What comes next

In Stage 6, CarryWell must earn its first real payments. The founders confirm a licensed route with the university, order a first batch they can afford, and sell to twenty named prospects (societies and individual students) with every message written and sent by themselves. To pass G6 they need C5 and C6 supported with at least moderate confidence, and each founder must redesign one real workflow around AI and defend it without AI help. For CarryWell that is likely to be stock and orders: the AI proposes a reorder from sales and stock levels; a founder approves every purchase.

What this example shows

1. **The first idea was not the business.** CarryWell started as an eco product. Twelve students who said waste "matters" and almost none who acted on it moved the founders to a problem people actually had: leaks and the microwave queue. Naventros treats that change as success, not failure.
2. **Online data shaped the questions, not the answers.** Search interest, marketplace listings and reviews gave the founders language and hypotheses. Only observation and interviews told them what students on their campus do.
3. **A human made every formal decision.** The mentor's CONDITIONAL in Stage 2 stopped a weak, after-the-fact conclusion from becoming a GO, and named exactly what would fix it.
4. **AI made the work faster and was overruled when it should be.** It critiqued, coded, tagged and drafted. The founders rejected an invented market size, a suggested segment with no evidence, and a product claim their own test had disproved.
5. **The cheapest test came from the gap list.** A five-unit bench test costing 110 CU removed a false promise before any money was spent on stock.
6. **Local rules shaped the tests.** Because student founders may not take money before a licensed route exists, the demand test collected intent, and the paid test waits for Stage 6.

CarryWell, Layla, Omar, the campus, the interviews, the online research results and all figures are fictional and were written to illustrate the method. They are not evidence about any real market, product, platform or institution.