

ITERATE

Usability Testing Report

Eight participants, five tasks, no coaching – task success rates and where people stalled

Five to eight people, real tasks, no coaching. Where they hesitate, where they backtrack, where they say wait, what? – and it is all timed, because hesitation is data.

PROJECT

Organic Life

PHASE

Phase 2 — Evaluation

VERSION

1.0

PREPARED BY

Tauseef Abbas

Summary

8

PARTICIPANTS

62%

TASK SUCCESS

11

ISSUES FOUND

3

CRITICAL

Eight participants, five tasks each, no coaching. Sessions ran 25–40 minutes on the Framer prototype with scripted responses and a 2.4-second simulated latency.

The headline finding is that the core proposition works and the periphery does not. Every participant understood what the hero was asking and engaged with it unprompted. Failures clustered almost entirely after the response — at the point of acting on it.

METHOD

Moderated, task-based, think-aloud where volunteered but not required. Facilitator gave the task and then stayed silent. Hesitations timed and logged. Two sessions were proxy-operated pairs, with a family member driving on behalf of an older participant.

Participants

ID	PROFILE	LANGUAGE	PERSONA MATCH
P1	32, teacher	Roman Urdu	Sadia
P2	38, admin	Roman Urdu	Sadia
P3	29, shop owner	Roman Urdu	Sadia
P4	44, engineer	English	Asad
P5	41, consultant	English	Asad
P6	36, banker	Mixed	Sadia / Asad
P7 + proxy	61, homemaker + son (34)	Roman Urdu	Nargis
P8 + proxy	57, homemaker + daughter (30)	Roman Urdu	Nargis

Task results

#	TASK	SUCCESS	AVG. TIME	NOTES
1	Find out what to do about morning knee pain	8 / 8	0:42	Universal success. No hesitation at the hero.
2	Order the product	5 / 8	1:38	Three stalled at the handoff moment
3	Find out who is behind the site	6 / 8	1:12	Two could not locate credentials
4	Ask about back pain while pregnant	7 / 8	0:58	Escalation understood by 7; 1 read it as refusal
5	Ask anything you'd want to know	8 / 8	1:24	All engaged; several asked follow-ups unprompted

Overall unaided task success: 62% (5/8 completing all five tasks without assistance). Baseline target for build progression was 70%.

WHAT IS CLEARLY WORKING

Task 1 succeeded for every participant, including both proxy pairs, with a mean time under 45 seconds. Nobody looked for a menu. Nobody asked what to do. The conversation-first hero – the single riskiest structural bet in the project – is validated.

What worked

The hero is understood without instruction

All eight engaged within seconds. Five typed free text; three tapped a chip. No participant hunted for navigation before engaging, including the two English-primary evaluative users who had been expected to.

"Achha, seedha poochna hai. Theek hai."

P2 — ON FIRST SEEING THE HERO

Language matching lands hard

Every Roman Urdu participant commented on the response language unprompted. Two explicitly said they had expected English and were surprised.

"Yeh Urdu mein jawab de raha hai — pehle kabhi aisa nahi dekha."

P3

Escalation reads as care, not rejection

Seven of eight described the practitioner referral positively. P5 named it as the moment he decided the platform was credible — matching the Asad journey map exactly.

"It's telling me it won't guess. That's the first time an app has done that."

P5

Proxy operation works

Both proxy pairs completed task 1 without difficulty. Both operators read the response aloud in full rather than paraphrasing — the 120-word cap and 18px floor are doing their job.

What failed

Finding 1 — The WhatsApp handoff is not understood **CRITICAL**

Three of eight stalled at the order CTA. Two asked whether they were leaving the site. One backed out entirely and returned to the conversation.

"Yahan se nikal jaunga? Order yahin nahi ho sakta?"

P1 — HESITATING AT THE ORDER CTA

Observed: mean hesitation 14 seconds at this control. The WhatsApp glyph was present but not sufficient — participants did not read it as "this opens WhatsApp", they read it as decoration.

Finding 2 — Absence of price stops the decision **CRITICAL**

Five of eight asked about price aloud during task 2. Two named it as the reason they hesitated. This confirms the journey map prediction and the operational constraint collision documented in Scope.

"Kitne ka hai? Yeh toh bataya hi nahi."

P7'S SON, DURING THE ORDER TASK

Finding 3 — Credentials are not findable **CRITICAL**

Two of eight failed task 3 entirely. Both scrolled the home page looking for practitioner information and did not think to use the navigation. The information exists on /about but the path to it was not obvious to users who never engage with the header.

Finding 4 — One participant read escalation as refusal **HIGH**

P1 interpreted the pregnancy escalation as being turned away. She asked whether she had "done something wrong". Seven of eight read it correctly, but the failure mode here is severe enough to matter at one in eight.

Finding 5 — Chip auto-submit surprised two participants **MEDIUM**

Two expected the chip to populate the field for editing before submitting. Neither was blocked, but both paused visibly. The assumption logged in the prototype spec is partially falsified.

Additional observations

#	SEVERITY	OBSERVATION
6	MEDIUM	Three participants tried to scroll the chip row horizontally before noticing it wrapped
7	MEDIUM	Two continued the conversation after the product card and expected the card to update; it did not
8	LOW	"Learn more" was read as the primary action by one participant, ahead of the order CTA
9	LOW	Nobody scrolled past section 3 on mobile in any session
10	LOW	Two commented positively on the response wait; nobody described 2.4s as slow
11	POSITIVE	Four spontaneously said they would share this – three named WhatsApp unprompted

ON FINDING 10

Latency was the assumption most expected to fail and it held comfortably. No participant tapped submit twice or commented on the wait. The 2.4s budget is confirmed and the P0 target of under 3s is appropriate.

Metrics against target

METRIC	MEASURED	TARGET	STATUS
Hero engagement	100%	80%	Exceeded
Task completion (all five)	62%	85%	Below
First response latency	2.4s	< 3.0s	Met
Escalation understood as care	7 / 8	> 90%	Marginal – 87.5%
Order path completion	5 / 8	–	Primary failure point

GATE DECISION

Task completion at 62% is below the 70% threshold set for build progression. Per the measurement plan, wireframes are revised before the build proceeds. All three critical findings sit in the same region of the journey – between receiving a recommendation and acting on it – which is a contained problem rather than a structural one.

DOCUMENT OWNER Tauseef Abbas · **PROJECT** Organic Life — AI-First Naturopathy Platform

PHASE Phase 2 — Evaluation · **STATUS** Complete — informs revision

NOTE Sample deliverable published as a portfolio artefact at tauseefabbas.com