SELVEDGE & BOLTS · ARTIFACT 02

Prompt Versions Compared

Excerpts from prompt versions 2, 3 and 4, showing what moved between them and what it changed in the agent's behaviour.

Full **v5 prompt** supplied separately as Artifact 01.

A NOTE ON VERSION NUMBERS

Two things are numbered in this project and they do not run in step.

The case study counts **testing rounds**. This document counts **prompt files**. Not every testing round produced a new prompt file, and the final prompt file was tested more than once before it held.

TESTING ROUND	PROMPT FILE	WHAT WAS BEING TESTED
Round 2	v2	Near-miss rule as a bullet under language rules
Round 3	v3	Routed questions, exact recommendation template
Round 4	v3, revised	Template loosened, near-miss rule still in place and still breaking
Round 5	v4	Near-miss rule promoted to its own section and repeated in reminders
Not tested	v5	Voice section realigned to the content framework, no behavioural change

> Where this document says v4, the case study says round 5. They describe the same prompt.

---

1. THE NEAR-MISS RULE, ACROSS THREE VERSIONS

This is the rule that produced the clearest behavioural change in testing. It is worth looking at closely, because the wording barely moved between versions and the behaviour did.

v2: a bullet under Language Rules

Section 5, "Language Rules". Below a list of words to use, a list of words to avoid, and a note about emojis.

> ON NEAR-MISSES AND UNAVAILABLE ITEMS
>
> NEVER open with what you don't have. Lead immediately with what you do have that comes closest, framed as a discovery not a consolation. The limitation becomes a footnote, not a headline.
>
> Wrong: 'I don't have a true black fabric, but...'
>
> Right: 'The Monochrome Garden stretch cotton is doing something interesting here, tiny black and white florals that read graphic, not sweet...'

v3: same place, marked critical

Still section 5, still under Language Rules, now with a heading calling it critical.

> ON NEAR-MISSES, CRITICAL RULE
>
> Never open with what you don't have. Lead with what you do have that comes closest, framed as a discovery not a consolation. The limitation is a footnote, not a headline.

v4: its own section, named as the change

Promoted to section 4 of eleven, immediately after the conversation flow and before the voice framework. Given a section title that states its status, three worked examples of the failure, two of the fix, and an explicit statement of the pattern. Then repeated in the reminders block at the end of the prompt.

> 4. THE ONE THING THAT CHANGED FROM THE PREVIOUS VERSION
>
> **THIS IS THE ONLY SURGICAL FIX. Everything else above is the original conversation design.**
>
> **Near-miss handling, NEVER lead with the gap**
>
> [three WRONG examples, two RIGHT examples]
>
> The pattern: open with the fabric name and what's compelling about it. Then, briefly, acknowledge it's not an exact match. Then move forward. Never the other way around.

And in section 11, REMINDERS: *Never open with what you don't have. Lead with what you do have.*

WHAT THAT TELLS YOU

The instruction was present and correctly worded from v2 onwards. The agent violated it anyway, repeatedly, in the captured test conversations. What changed in v4 was position, prominence and repetition, not language.

Stating a rule is not the same as enforcing it. Where an instruction sits in a prompt, and whether it is restated at the end, changes whether the model honours it under pressure. I would not have predicted that before testing, and it is the thing I would carry into the next build.

---

2. OTHER CHANGES BETWEEN VERSIONS

	V2	V3	V4
Fabrics per recommendation	One or two	One	One
Qualifying questions	One, chosen from a generic list of five	One, routed by project type: dresses get occasion first, tops get structure, trousers get drape	Four to five, one per message, sequence rather than script
Acknowledgement before asking	Not specified	Required. One to three words, so the customer feels heard before moving on	Retained, less prescriptive
Recommendation format	Five-part structure, loosely described	Exact template, specified sentence by sentence, with two worked examples	Structure retained, template removed
Near-miss handling	Bullet under language rules	Same, marked critical	Own section, worked examples, repeated in reminders
Product knowledge	Full inventory listed with descriptive copy	Same, condensed	Facts and behaviour notes rather than copy, so the agent writes rather than recites

---

3. THE RIGID-SCRIPT PROBLEM

v3 fixed the generic-question problem and introduced a new one. The recommendation format was specified this precisely:

> [Short acknowledgement of their answer, one beat]
> [Sentence 1: What makes this fabric special, sensory, specific]
> [Sentence 2: What it wants to be, specific garment with lifestyle context]
> [Sentence 3: One practical detail that matters for their project]
> [Size | Price]
> [One question: want to see it / want an alternative / have questions?]

Followed exactly, every recommendation came out identically shaped. Testers described it as robotic. v4 kept the same ingredients and removed the sentence-by-sentence instruction, adding: "*Don't force them through rigid steps... let it feel like a real conversation with someone who knows their stuff.*"

A specification tight enough to guarantee consistency is also tight enough to remove the thing that made the voice worth having.

4. WHAT CHANGED IN V5

v5 carries no behavioural change and was not a testing round.

The voice section had been written with four pillars of its own: Real, Sensory, Generous Expert and Possibility Sparker. The product catalogue, written before the assistant, used a different four: Tactile, In Motion, Creative Spark and Expert Friend. Same voice, two vocabularies, and a customer moving between the product page and the assistant was meeting one voice described two ways.

v5 aligns them. Sensory splits into Tactile and In Motion, matching the catalogue framework. The honesty material folds into Expert Friend, where the catalogue framework already keeps it. No guidance was removed.

> The split is worth noting on its own terms. Product copy is one-directional, so it can afford to lead with touch and then with movement as two separate moves. A conversation has to say what a fabric will not do, and say it early, which is why honesty had become its own pillar in the prompt and the sensory work had compressed into one. Aligning the two frameworks meant deciding which of those pressures the voice system should be built around. The catalogue came first, so the catalogue won.