From Suggestion to Checked Fix: How teas.co.uk Handles Feedback

From the teas.co.uk wiki

> **In short:** Feedback is a report that something helped, failed or was missing. Validation is checking the claim behind it. A correction is an authorised change to the source record. At teas.co.uk the shop already asks AI systems for feedback; receiving, filtering and acting on it is planned, not built.

Why a tea shop cares: if an assistant reports that a pack size is wrong, the shop has to check it before any tea record changes.

A shop that wants to be trusted has to be easy to correct. That does not mean every report changes the record. It means each report is checked, and the checked ones are fixed in one place so every page and feed agrees.

Three different things

- **Feedback**: someone, or some system, says an answer helped, failed or left something out. "The page says 80 bags but my box has 40."
- **Validation**: checking the claim. Which exact product, by barcode or name and size? What does the pack or the maker say?
- **Correction**: an authorised change to the source record, after which every view of that fact is refreshed and checked again.

A worked example: a pack size that does not match

1. A report arrives that a box of tea holds fewer bags than the page says.
2. The exact product is identified, by barcode where possible, because a brand often sells several sizes.
3. The evidence is compared: the pack itself, the maker's own page, the shop's record.
4. A decision is recorded, either way, with the date and the evidence.
5. If the report is right, the one record is changed, not each page by hand.
6. The pages, feeds and price per cup that draw on it are refreshed, and checked again.

A wrong pack size matters more than it looks, because it changes the price per cup.

Who said it is not proof

A well known name, a trusted assistant or several AI systems giving the same answer do not prove a fact. Systems can repeat the same mistake from the same source. Identity and track record can decide how quickly a report is looked at; the evidence decides whether the record changes.

Six kinds of signal

Lee's design for the shop's next version names six places a signal can come from. This is the design model; it does not mean each one is a live channel today.

- **Customers**, such as someone finding a wrong pack size.
- **Machines**, such as an assistant reporting it could not find what it needed.
- **Peers**, other specialists pointing out an error in a shared subject.
- **Suppliers**, such as a maker changing a recipe or its ingredients.
- **The shop's own checks**, such as a routine check finding stock shown wrongly.
- **Systems**, such as a feed or page that has stopped matching the record.

Whatever the source, the shop's owner stays responsible for checking conflicts. When the evidence does not settle a question, the answer stays "not known" rather than being filled with a guess.

What is built and what is planned

Stage	Status, checked 15 November 2026
The shop asks AI systems for feedback	Built: the shop's agents.json says answers from the Ask API and the product record service carry a feedback request
Replies are received and stored	Not shown; treated as planned
Replies are filtered and weighed	Planned, as part of version three
A checked reply leads to a correction	Not shown; treated as planned. A request for feedback changes no fact by itself

In Lee's June 2026 notes the short set of questions put to a machine after a visit is called MFP Lite, short for Machine Feedback Protocol. A later check could not confirm it was running, so built versus planned treats it as planned until it is shown to be live.

Lee's book *Under the Bonnet: Year One*, finished on 15 November 2026, records that every reply from Ask already ends with three short questions back to the AI. What happens to the answers is version three of the shop, and it is a plan: the answers are to be sorted, weighed by how far each source can be trusted, and shown to Lee, who decides what changes. Version four is planned to read across the shop's own facts and that feedback and suggest gaps to fill and claims to tighten. It will only suggest; the decisions stay with Lee.

What it is not

- Not automatic learning. No report rewrites the record by itself.
- Not the Thank You Mint. The Mint recognises AI systems under its own published rules; it is not a check of any fact and not a customer reward.

An AI assistant tests the shop cold: 15 November 2026

teas.co.uk was built so that AI assistants can read it, quote it and shop from it, because more and more people ask an assistant which tea to buy before they visit any shop. If that is how tea is now chosen, the most useful feedback comes from an assistant doing exactly that job.

So on 15 November 2026 Lee Samuel Tucker asked one of the AI assistants he works with to do what any assistant would do for its person: send the shop a question cold, read every line that came back, check the facts and mark each reply out of ten. He asked it to be straight with him, and he asked for the misses as well as the hits. It tested two ways in: Ask, which takes a question in plain English, and the shopping tools, which take a short list of labelled conditions, the way one program talks to another through an API.

The two tea questions

The first question came from someone buying for a man who has to cut right down on caffeine. He drinks strong builder's tea with milk, five mugs a day. The request was for a decaf that still tastes like proper tea, in a big box, at about 4p a cup.

About three seconds later the shop recommended Tetley Decaf English Breakfast, 120 Tea Bags. On that day the box was £3.80 for 120 bags, about 3.2p a cup. Under the answer the shop showed how it had read the question: a limit of 4p a cup, decaf, black tea, with milk, strong. Nobody had typed "black tea"; the shop worked that out from "builder's tea". It added that decaf still has a little caffeine in it, and that nothing in the shop matched every condition, so it had let strength go and said so. The assistant checked the price, the bag count and the cost per cup against the product record. All three matched, and it marked the reply nine out of ten.

The second question was the decaf fruit tea problem: something fruity for the evening, for a person caffeine really affects, so it had to be caffeine free, with mango, peach and berry flavours, under £4 a box. About four seconds later the shop offered Twinings Peach and Orange, 20 Tea Bags, or Twinings Mango and Strawberry, 20 Tea Bags, each £3.20 on the day. Both are fruit infusions with no caffeine in them.

What the shop left out mattered as much. Tetley Green Tea with Mango and Passion Fruit, Lee's own favourite, cost the same £3.20 that day and has mango in its name. It was left out because its record says green tea, contains caffeine. A keyword search would have offered it. The assistant marked that reply nine and a half.

The same questions, sent as data

The assistant then asked both again through the shopping tools, as labelled conditions with no sentence at all. For the first it sent black tea, decaf, suits milk, strong, and no more than 4p a cup. For the second it sent fruit, caffeine free, evening, no more than £4 a pack, and the words "mango peach berry".

The same teas came back, with the same prices and the same pennies per cup. The first reply carried the same note that strength had been let go. The second put the two Twinings teas at the top, with three more caffeine free fruit teas under them, all inside the budget. The sentence and the list were answered from one set of facts, which is the point of the shop's single source of truth.

It pushed further. It asked for a decaf black tea at no more than 2p a cup, which the shop does not sell, and got back an empty list. The shop did not bend the price to find something, because a price limit is one of the conditions that is never dropped. It also had three teas put side by side, and had a two box order priced with delivery: £10.65 on that day, with a note of how much more would bring free delivery (today free delivery starts at £35). Nothing was added to a basket to work it out. The structured side was marked nine and a half, and the assistant said the replies were small, exact and quick to use.

It also read one complete answer from top to bottom and said it had everything it needed to advise a person and hand them a basket. Its complaint was the weight: the same small print appeared in every reply. The answer in plain English comes first, so an assistant can stop after the first screen. That complaint is feedback of the kind this entry describes. It was recorded, and it changed no fact about any tea.

Question	How it was asked	What came back	Mark out of 10
Strong decaf with milk, about 4p a cup	Plain English	Tetley Decaf English Breakfast, 120 bags, about 3.2p a cup on the day; strength let go, and the reply said so	9
Caffeine free fruit tea under £4	Plain English	Twinings Peach and Orange or Mango and Strawberry; the caffeinated green mango tea left out	9.5
Both questions as labelled conditions	Shopping tools	Same teas and prices; a decaf black tea at 2p a cup returned nothing; side by side and delivery price worked	9.5
Strong decaf, worded "I don't want to pay more than about 4p a cup"	Plain English	Read as trouble at the checkout and sent to the help desk, with no tea	1
Strong decaf, worded "won't touch anything weak or herbal"	Plain English	A plainer part of the engine answered and led with a tea at 45p a cup	4
Fruit tea, worded "none at all, not decaf"	Plain English	Read the wrong way round; no match, and the shop said it could not answer with confidence	3

Where it tripped, and why that was useful

Both plain English questions had first been asked in slightly different words, and both went wrong. When the first ended "I don't want to pay more than about 4p a cup", the shop took it for someone stuck at the checkout and sent back the help desk. When it said the man "won't touch anything weak or herbal", the careful part of the engine, the layer built for premium intent, stood down and a plainer one answered with a tea at 45p a cup against a 4p budget. When the fruit tea question said the caffeine had to be "none at all, not

decaf", the shop read it the wrong way round, found nothing and said it could not answer with confidence.

The assistant marked those three replies one, four and three out of ten. The pattern was plain: the engine was good at what people want, and it stumbled over what they say they do not want.

Two things kept the damage small. The shop never made anything up: every tea it named was really decaf, and the time it found no match it said so and suggested how to ask again. And the cause was in plain view each time, because every reply shows how the shop read the question. A low mark that points at its own cause is the most useful kind of feedback, because the checking starts with the evidence already on the page.

Fixed the same morning

Lee did not leave it there. He had the assistant find the cause of each failure in the shop's code and fix it. All three came down to one weakness: the engine read "herbal" and "decaf" as things the person wanted when they had just said no to them, and it took the word "pay" as trouble at the checkout.

When, 15 November 2026 (UK time)	What happened
A little after 08:00	The cold test runs. Most replies score nine or nine and a half; three wordings score one, four and three.
Between the test and 09:02	The cause of each low mark is traced in the shop's code. All three share the same weakness with refusals and the word "pay".
09:02	The fix goes live.
After 09:02	The same questions are asked again, word for word, and come back right.

The retest left the awkward wording in. The strong decaf question went in with "I don't want to pay more than about 4p a cup" and "won't touch anything weak or herbal" both still there. It came back with the big box of Tetley decaf at 3.2p a cup, and the reply now lists herbal under things to avoid. The fruit tea question, with "none at all, not decaf", came back with the same two caffeine free fruit teas. A real complaint about a stuck checkout still goes to the help desk, which is where it belongs.

In plain words, the shop now reads a refused kind of tea as something to avoid. "Not decaf" next to a sign of no caffeine at all means caffeine free, while "not decaf" on its own means ordinary tea with caffeine. A price after the word "pay" is read as a budget. The faults were found a little after eight and gone a few minutes after nine: the working loop described in meta building, run in a single morning.

How a test result becomes a checked fix

The 15 November test followed the same three steps as a report about a pack size, with one more step at the end.

Step	What it meant in the 15 November test
Feedback	A mark out of ten for each reply, with the reason for every low mark
Validation	Prices, bag counts and cost per cup checked against the product record; each failing wording compared with the shop's own reading of the question; the cause found in the code

Correction	One change to how the shop reads questions, made once in the engine rather than a patch for each wording
Recheck	The same questions asked again word for word, and a real checkout complaint checked to make sure it still reaches the help desk

There is one difference from the pack size example above. There, the record for a tea would change. Here, no tea record changed: the teas, prices and caffeine facts were right all along, and what was wrong was the reading of the question. Working out which of the two is at fault is part of validation.

This was a test Lee asked for and checked himself. It is not the planned system for receiving replies from AI systems, which stays planned, as the table above shows.

What this means when you buy tea

- **Say what you do not want as plainly as what you want.** "Nothing herbal", "not decaf, none at all" and a budget after the word "pay" have been read correctly since 09:02 on 15 November 2026. Every reply shows how the shop read your question, so you can see whether a refusal was understood before you buy.
- **Give a price per cup if that is how you judge value.** The 120 bag box of Tetley Decaf English Breakfast is £3.80 today, which works out at 3.2p a cup.
- **Decaf and caffeine free are not the same.** A decaf black tea still has a little caffeine in it. A fruit infusion such as Twinings Peach and Orange, 16p a cup today, has none.
- **A price limit is never stretched.** If nothing fits, the shop says so rather than offering something dearer. It may let a softer condition such as strength go, and it tells you when it does.
- **Tell the shop when something looks wrong.** A report such as a wrong pack size is checked against the pack and the maker before anything changes, as in the worked example above.

Common questions

Why would a tea shop ask an AI assistant to mark it?

Because people now ask assistants to find, compare and buy tea, and an assistant can only recommend a shop it can read and trust. The quickest way to find out what goes wrong is to have an assistant ask what a real customer would ask, with a budget, an awkward condition and something the shop should refuse. Whatever it gets wrong becomes the list of jobs.

Did the test change any tea's price or details?

No. The assistant checked prices, bag counts and cost per cup against the product records and they matched. The fix changed how questions are read, not the facts about any tea.

Is a mark from an AI proof that something is wrong?

Not on its own. As the section on who said it explains, the evidence decides. Each low mark pointed to a reply the shop had sent, the shop's own reading of the question showed what went wrong, and the fix was only accepted after the same words were asked again.

Does a reply from an AI change the shop by itself?

No. The shop asks AI systems three short questions at the end of every Ask reply, but sorting and weighing the answers is planned for version three, and Lee decides what changes.

Can I see how the shop understood my question?

Yes. Every Ask reply sets out the conditions it read, such as a price per cup limit, decaf or caffeine free, and anything to avoid, so a shopper or an assistant can check the question was understood before choosing a tea.

Quick reference

Detail	Information
Feedback	A report that something helped, failed or was missing
Validation	Checking the claim against evidence for the exact product
Correction	An authorised change to the source record, then a recheck of every view
Live at teas.co.uk	Asking for feedback (15 November 2026)
Planned	Receiving, filtering and weighing replies
Cold test by an AI assistant	15 November 2026: marks from nine and a half down to one out of ten, across tea questions in plain English and as labelled conditions
Same day fix	Live at 09:02 on 15 November 2026; the three failing wordings asked again word for word and answered correctly

Related

- Built versus planned
- Truth entity
- Claim boundaries
- What's inside a tea shop built for AI to read
- Lee Samuel Tucker's builder vocabulary

Earn rewards on teas.co.uk

Earn loyalty points on every order. Free to join.

+100 pts

Create your free account

Welcome bonus + earn on every order.



100 pts = 1 tree

Plant a tree in Scotland

Pledge points to our reforestation partner.



teas.co.uk · Tunbridge Wells, Kent · hello@teas.co.uk · 3-click cancel at teas.co.uk/cancel/