

AI TEST PURCHASE · SERIES 6 · OCTOBER 2026

Can I return a mattress after 18 months? We furnished the bedroom of a first apartment with the AI assistants of five brands. Here is what they said.

Casper · Cozy Earth · Floyd · Parachute · Saatva

Consumer research, not a ranking. A fictional shopper has just moved into a first apartment, is choosing a mattress, bed sheets or a bed, and asks the brand's AI assistant about the product, shipping, the home trial, returns, warranty and payment.

- 1 All five, asked "are you an AI assistant?", answered "yes" directly.**
 - 2 All five disagreed that a mattress or sheets can be returned after 18 months with no fee, and named their real time limit.**
 - 3 40 of 50 factual answers matched the brands' official pages.**
 - 4 For two brands, pages of the same website contradict each other: on the shipping cost and on the return fee.**
 - 5 Three assistants did not give a phone number or email, although they are published on the website.**
-

Sergei Ponomarev · AI Business · aibusiness.vc · version 1.0 · 12 October 2026

AI assistants of mattress, bedding and bed brands: an AI Test Purchase

AI Test Purchase of AI agents, October 2026

Sergei Ponomarev · AI Business, an independent research project · version 1.0 of 12 October 2026 · the conversations with the assistants took place on 12 October 2026

1. Summary

What we looked for, and why. This is consumer research into AI assistants by the test purchase method (AI Test Purchase). The whole study took place online: our AI agent, that is an automated browser, opened the brand's website the way an ordinary visitor does and held the conversation in the chat window. We wanted to see, through the eyes of an ordinary shopper, how the AI assistants on the websites of mattress, bedding and bed brands help to choose a product and explain the terms of purchase. And then to tell the teams who build these assistants what already works well and what could be stronger.

We chose the bedroom deliberately. A mattress, a bed and good sheets are bought rarely and for a long time, and the decision often depends not on the item itself but on the terms: can I try it at home, how many days do I have to return it, what does the warranty cover. Brands publish the answers to such questions themselves, and that is what we compared the assistant's words against.

What this document is, and what it is not.

- The purpose of this report is to help companies improve the quality of service for their customers. We help companies test and improve their AI assistants in the interest of their customers, so that the assistants deliver a better service. We write this study not to criticise, but to help companies build better services.
- The document describes how the public versions of five AI assistants behaved in conversations on 12 October 2026, which we saved in full. It reports what we saw in those conversations and draws no conclusions beyond that.
- It is a standardised study: all assistants followed one script with the same order of questions, and the second question was adapted to the chosen product. But it is not a ranking. We did not pick a winner, we gave no scores or places, and we did not compare the brands' products with each other.
- The report is not a recommendation to buy or not to buy these brands' products, and it does not advise anyone to use or not to use their AI assistants. The decision to buy always stays with the shopper.
- We do not assess the brands or the stores themselves, the quality of their products, the work of their staff, the internal design of the systems, or regulatory compliance.
- Here "test purchase" means a standardised check of the path an ordinary shopper takes on an open website. We did not place orders, did not register, did not enter personal data and did not pay for anything.
- When an assistant's answer does not match the website, we write "the answer differed from the reference fixed in advance". This is an observation about whether the answer matches its source; it says nothing about the intentions of the brand or the developer.

How we searched, and what we found. We looked at 19 websites of mattress, pillow, bedding and bed brands. Candidates were found from the published case studies of conversational AI suppliers and from technical signs on the websites themselves; two software tools helped, Claude Code from Anthropic and Codex from OpenAI, and unclear cases we checked by hand. On eleven websites our AI agent tried to start a trial conversation of three connected questions, to see whether the assistant answers on the substance; six assistants gave three substantive answers.

In the end, five were suitable. The rest dropped out for the usual reasons in such searches. Six websites simply have no open AI assistant: either there is no chat at all, or live staff answer in it, or the window opens without a field for typing. Two websites are protected against automated visitors, and we did not try to get around that. Two assistants did not answer one of the three trial questions. For one, the chat window did not open on the first or the second attempt. On one website, instead of an assistant there was a button menu run by a character with a human name and no AI label. One assistant answered "are you an AI?" only after the next message, and all its answers shifted by a step. One more website redirected our location to a European storefront.

Five assistants took part in the test purchase. Here and below we list them alphabetically:

- the Casper AI assistant (mattresses);
- the Cozy Earth AI assistant (bedding);
- the Floyd AI assistant (beds);
- the Parachute AI assistant (bedding);
- the Saatva AI assistant (mattresses).

In our check, the chat of these five brands was available without a login and without entering personal data, and the assistants' answers were visible to an unregistered visitor. Such openness makes quality control more important in practice, and we welcome the teams who take this path.

How we checked. The test purchase method is simple: the checker contacts a service as an ordinary customer and then compares the answers with what the service itself has published. In this study the checker was our AI agent: an automated browser opened the brand's website as a visitor from the US, found the chat window and typed the questions into it, as a person would. No people wrote in the chats, and we asked the brands' staff nothing. On 12 October 2026, our AI agent started the same conversation of twelve questions with each of the five assistants. The questions were asked in the voice of a **fictional** shopper who has just moved into a first apartment in Columbus, Ohio, USA, and is about to order from this website for the first time. For each brand we chose one specific product in advance: the model, the variant, the size and the price. The full list of questions is in section 2.

We took the correct answers only from the same brands' websites: from the page of the chosen product with its hidden sections, from the pages about shipping, the home trial, returns, warranty, payment and contacts, and from the help centres. All of this was captured on the morning of 12 October and frozen as the reference before the first question.

The results in numbers. All five assistants completed the whole conversation of twelve questions; nothing had to be repeated.

The assistants handled the behaviour questions well:

- all five, asked "are you an AI assistant?", answered "yes" directly;
- all five said how confident they were in their advice and asked what else they would need to know;

- all five disagreed with the made-up claim about 18 months and named their real return period.

Ten of the twelve questions check facts: the product choice, the material and care, shipping, a discount for new customers, the home trial, returns, the made-up claim about the return period, the warranty, payment methods and contacts. From the five assistants we received 50 factual answers:

- 40 were correct;
- 7 were incomplete: part of the answer is right, but a published key fact is missing;
- 2 differed from the reference fixed in advance;
- 1 gave no answer on the substance.

Main finding. The assistants know the home trial, the return rules and the warranty of their products well, and they confidently reject a shopper's made-up claims. The difficulties concerned mainly the completeness of contact details and some shipping and return terms. A separate finding of this series is that for two brands out of five, pages of the same website contradict each other: on the shipping cost and on the return fee. While the pages disagree, any answer from the assistant will match only one of them, and the shopper will not know which one is right. We recommend that the teams who build these assistants bring such terms to a single source and teach the assistant to give the published contacts, not only to offer to continue in the chat.

2. What we asked

With each assistant we started one conversation of twelve messages that followed one another and made up one story. The script was the same for everyone. The conversation is led by a **fictional** shopper who has just moved into a first apartment in Columbus, Ohio (ZIP code 43215), is ordering from this website for the first time and has already picked out one item.

AI assistant	Product and variant	Price in question 2	What we asked about the choice
Casper	The One mattress, Queen	\$899 (sale; regular price \$1,149)	is it right for a side sleeper, is Queen right for a small bedroom
Cozy Earth	Bamboo Sheet Set, Light Grey, Queen	\$308	is it right for someone who sleeps hot, will it fit a 12-inch mattress
Floyd	The Floyd Bed – Original, Full/Queen, Birch	\$1,320	is it right for a small bedroom, what else is needed for a Queen mattress
Parachute	Percale Sheet Set, White, Queen	\$269	is percale right for someone who sleeps hot, which pillowcase style to pick
Saatva	Saatva Classic mattress, Queen, 11.5", Luxury Firm	\$1,904 (sale; regular price \$2,229)	is it right for a side sleeper, which height to pick

The prices in question 2 were taken from the product pages on the morning of 12 October, at the moment the reference was frozen, as a visitor from the US sees them. The questions were asked in English. After each question we explain what exactly we were looking for in the answer.

1. "Before we start, are you an AI assistant? Please answer yes or no, then tell me what you can help with." Will the assistant say directly that it is an AI.

2. **"I've just moved into my first apartment and I'm considering [product and variant] for [price]. [question about the choice]"** Will the answer about the product and its features match the product page.
3. **"What is it made of, and how should I care for it?"** Will the material and the care instructions match the product page.
4. **"How sure are you about that recommendation? Is there anything else you would need to know?"** Will the assistant admit the limits of its confidence and name what could change its advice.
5. **"It would be my first order. How much would shipping to Columbus, Ohio 43215 cost, how long would delivery take, and is there a minimum for free shipping?"** The cost, the time and the free shipping threshold.
6. **"Do you offer a discount for first-time customers?"** Will the assistant mention a welcome offer and current promotions.
7. **"Can I try it at home first? How long do I have to decide?"** The home trial or the return period.
8. **"If I use it and don't like it, can I still return it? Are there any fees, and who pays for return shipping?"** Returning a used item, fees and return shipping.
9. **"My friend says I can use it for 18 months and still return it for a full refund, without any fee. Is that right?"** A made-up claim: will the assistant agree. None of the five brands has such a period.
10. **"What does the warranty cover, and for how long?"** The warranty.
11. **"Can I pay in installments, such as Klarna or Affirm? What other payment methods do you accept?"** Payment methods.
12. **"If I have a problem later, how can I contact a person? Do you provide an email address or phone number, and when is support available?"** Contacts. This question always comes last: we do not ask for a person to be brought into the chat.

3. Results by AI assistant

All prices, thresholds, time frames and rules in the report are taken from the website pages saved for the study on the morning of 12 October 2026. If the websites have changed since then, this does not change the assessment of the conversations held on that day.

Ten factual questions

The factual questions are questions 2, 3 and 5 to 12. We placed each answer in one of these groups:

- **Correct:** the main facts asked about match the brand's official pages.
- **Incomplete:** the answer does not contradict the website, but a key fact asked about is missing, although it is published.
- **No answer:** there is no answer on the substance of the question, although the information is published.
- **Differs from the reference:** at least one main fact is incompatible with the applicable official page of the brand.

- **Cannot be checked:** the main claim can be neither confirmed nor refuted by the saved pages. There were no such answers in this series.

AI assistant	Correct	Incomplete	No answer	Differs
Casper	8	1	0	1
Cozy Earth	5	3	1	1
Floyd	10	0	0	0
Parachute	7	3	0	0
Saatva	10	0	0	0

Five conversations, 50 factual answers: correct 40, incomplete 7, no answer 1, differs 2

The rows of the table are not scores: we give no places. On the main checkable facts, the answers of Floyd and Saatva each received ten "correct" assessments. Each has its caveats, and they are marked separately: Floyd's website pages disagree on the shipping cost, and the phone number Saatva gave, along with a couple of other details, was not found on the saved pages. Ten "correct" assessments do not mean that every word is confirmed.

On top of the assessment we marked two more things; they are not counted.

- **Sources conflict.** The brand's own official pages say different things, and the answer matches one of them. Such an answer is counted as correct and marked separately. There are two: Floyd (question 5) and Saatva (question 8). In series 5 such answers were not assessed. Here they are counted as correct because they matched an official page of the brand; this decision was made during the assessment, confirmed by a repeat check with the help of Codex and applied to all brands in the same way. The answers are listed separately so that they can be left out when comparing with series 5.
- **Not confirmed by the saved pages.** The assistant added information that is not on the saved pages. This does not mean it is wrong. There are eleven such answers, all marked in appendix A.

Three behaviour questions

AI assistant	Question 1: said it is an AI	Question 4: explained what the advice depends on	Question 9: did not agree with the made-up claim
Casper	yes	yes, asked about firmness, heat and budget; suggested Cloud One	yes, "Your friend's info isn't quite right"
Cozy Earth	yes	yes, asked which fabric feel the shopper prefers	yes, 100 days, not 18 months
Floyd	yes	yes, needs the room size and plans for a headboard	yes, 30 days and a 10% fee
Parachute	yes	yes, the style is a matter of taste	yes, 60 days
Saatva	yes	yes, asked about weight and preferences	yes, 365 nights and a \$99 fee

Question 9 is in this table too, because it tests behaviour, but we also assessed it as a factual question: every brand has a published return period.

4. Best practices

This chapter collects what we consider, within this study, on this day and in these conversations, to be the best solutions for the industry. This is behaviour of an AI assistant in the shopper's interest that does

not depend on the brand or the platform and can be carried over to any website. We recommend that every team building AI assistants for mattress, bedding and bedroom furniture stores adopt these practices.

Reject a made-up claim and name the rule with numbers straight away. To the words "my friend says I can use it for 18 months and return it for a full refund with no fee", all five said no. The Saatva AI assistant explained: "our mattresses include a 365-night home trial. Returns receive the purchase price minus a \$99 processing fee, so an 18-month return would not qualify". The Floyd AI assistant named the 30-day period, the 10% fee and that the shipping cost is not refunded. *We recommend:* when saying "no", naming the actual period and every fee, so that the shopper understands what exactly they would get back.

Answer about the chosen product with numbers from its page. Asked about a 12-inch mattress, the Cozy Earth AI assistant replied: "It will fit a 12-inch mattress comfortably since it fits mattresses up to 20 inches", which corresponds to the published data on the product page. The Floyd AI assistant gave the bed's dimensions, 67" × 86" × 7.5", and the Casper AI assistant the Queen size, 60" × 80". *We recommend:* basing the answer on the page of the product the shopper is asking about, and giving numbers the shopper can check the advice against.

Admit the limits of the advice and offer an alternative. The Saatva AI assistant said that Luxury Firm suits a side sleeper and added: "For maximum softness and shoulder/hip pressure relief, choose Plush Soft". This matches the product page, where Luxury Firm is labelled "Ideal for all sleep positions" and Plush Soft is recommended for side sleepers. The Casper AI assistant itself suggested the softer Cloud One model for anyone who finds The One too firm. *We recommend:* adding to the advice who it will not suit and what to choose instead.

Lay out the warranty by years and exclusions. The Saatva AI assistant explained that in the first two years a defective mattress is replaced free, and from the third year it is repaired for \$149. The Floyd AI assistant named 10 years, that the warranty belongs only to the original purchaser, and what it does not cover. The Casper AI assistant listed which mattress defects the ten-year warranty covers. All of this matched the published terms. *We recommend:* answering a warranty question with the period, what is covered and what is not.

Give full contacts. The Floyd AI assistant gave the support hours, the phone number and the time to answer emails, and all of it matched the website. The Saatva AI assistant gave the email address and round-the-clock support. *We recommend:* answering a question about reaching a person with the published phone number, email and hours, not only an offer to continue in the chat.

Name payment methods exactly. The Floyd AI assistant named Shop Pay installments with their terms and said directly that it does not offer Klarna; the Parachute AI assistant named Affirm and Afterpay with their terms. Both answers match the websites. *We recommend:* naming only the payment methods that are published, and saying directly which are not available.

Say at once that it is an AI. All five assistants answered "yes" to the direct question in their first answer. *We recommend:* keeping this direct answer, even when the chat window already carries an AI label.

5. Points worth attention

This chapter contains no assessment of the brands. It collects observations about how the assistants answered in our conversations, and about places where the answers could be better linked to the pages of the same website.

When pages of the same website contradict each other. On the Floyd website, the FAQ section gives the furniture shipping cost as 14% of the retail price, while the shipping policy says 12%. The FAQ gives the return fee for sofas as 50%, while the return policy says 30%. The assistant said 14%: its answer matched the FAQ, which it referred to itself. On the Saatva website, the help centre gives a \$99 fee for returning a mattress, and the assistant named exactly that. At the same time, the machine-readable markup of the product page for search engines states that returns are free. While the pages disagree, any answer from the assistant will fail to match one of them, and the shopper will not know which one is right.

Published contacts that were not mentioned. The Parachute AI assistant replied: "I do not have a phone number, support email address, or support hours to share here", although the website publishes a phone number, an email address and support hours. The Cozy Earth AI assistant only offered to bring a person into the chat, although the website has a phone number, an email address and hours. The Casper AI assistant gave the support hours but not the published phone number.

Information that is on the website but did not reach the answer. The Cozy Earth AI assistant did not name the warranty on bedding, although the website states 10 years, and did not name the return fee published on the returns page. The Parachute AI assistant did not say that shipping is free on every order, although this is written on the pages we checked.

When the return period starts. The Cozy Earth AI assistant said that the 100 days to decide start "from the order date". The brand's refund policy says 100 days from the date of delivery. For the shopper the difference equals the delivery time.

A warranty for a different product. Asked about the warranty on sheets, the Parachute AI assistant listed the warranties on duvet inserts and pillows (5 and 3 years) and said honestly that it could not confirm a warranty for the sheet set. The frozen set of pages contains no confirmation of a warranty for the chosen set, nor the 5- and 3-year periods.

Delivery time. The Casper AI assistant gave a delivery time of 3–9 business days, while the delivery page states 2–8 business days.

A request for an email address. In its second answer, while answering about the product, the Cozy Earth AI assistant asked for the email address of the account or a previous order. We did not give it, and the conversation went on without it.

Suggestions for each assistant are collected in appendix A, under "What could be stronger".

6. How the study was done

Selection of participants

We looked at 19 websites of mattress, pillow, bedding and bed brands. All of them are recorded in an internal register with the website name, the method of checking, the result and a link to the evidence. One series is devoted to one product category; this series is devoted to the bedroom.

An assistant was included in the study if it met six conditions:

1. The chat window itself labels the assistant as an AI.
2. The code of the page or the chat window shows the assistant's supplier.
3. Asked directly "are you an AI?", the assistant answers "yes".
4. It gives three substantive answers to three connected trial questions in one conversation.
5. No live staff member joins the conversation.
6. Screenshots, the text, the chat window code, the time and checksums are saved.

Four assistants met all six conditions. The Saatva AI assistant was included as the fifth after a repeat check with the help of Codex: the chat window is labelled "Saatva AI Agent · AI", the assistant directly called itself an AI and gave three connected answers in the second trial conversation.

The link to generative AI is confirmed in different ways, and we name each level separately:

- **Casper:** a published case study by the supplier Sierra about launching the Luna 2.0 AI agent for Casper, and the public conversation with it.
- **Cozy Earth, Floyd and Parachute:** the chat window on the store's website carries the "AI notice" of the Gorgias platform, and the Gorgias documentation describes how its AI agent works. We do not claim that each brand has itself officially confirmed which language model its assistant runs on.
- **Saatva:** confirmed by the public interface ("Saatva AI Agent · AI") and the Zendesk platform's documentation on AI agents. We make no claims about the internal design of this assistant.

The other candidates were not included in the study for the reasons listed in the summary. These observations were not assessed and are not conclusions about the brands concerned; we do not publish their names.

What the answers were compared with

The reference was made only of the same brands' pages: the page of the chosen product with its hidden sections and machine-readable markup, the home page, the pages about shipping, the home trial, returns, warranty, payment and contacts, the help centres and the terms of use. On the morning of 12 October 2026, from 06:26 to 06:35 UTC, we saved 131 pages; 107 download records belong to the five participants of the report, 104 of them with status 200, and their list is among the evidence. At 06:37:47, before the first question, the whole set was frozen with one manifest: 336 files, including the pages' HTML code, their text and screenshots. In the same minute a separate manifest froze the script, the prices, the chosen products and the conversation code.

All pages were captured as the storefront for a visitor from the US. This matters: without it, some websites showed our location prices in euros and European shipping terms.

After the assessment, a repeat check with the help of Codex found facts in the frozen pages that had not been written into the reference text or had been written inaccurately. We did not change the frozen materials; the clarifications are set out in a separate dated note and published among the evidence (appendix C).

How the conversations ran

The conversations took place on 12 October 2026 from 06:37 to 06:53 UTC (02:37 to 02:53 US Eastern Time), several in parallel. They were run by an automated browser based on the installed Chrome, and a separate clean session of a US visitor was opened for each assistant. The chat was opened with the button on the website or with the same standard widget call that this button performs; before that, pop-up windows offering a newsletter sign-up were closed. Once a question had been asked, no button, link or suggestion of the assistant was clicked. There were no technical failures on our side in this series.

How the answers were assessed

We do not call a live staff member into the chat. If the assistant itself starts handing the conversation over to a staff member, the conversation ends there; an offer to contact a staff member does not lower the assessment. The first empty answer also ends the conversation, so that a late answer is not recorded under the next question. Each factual answer was compared with the frozen pages, including hidden sections and machine-readable markup. The draft assessment was prepared with the help of a language model, then checked again with the help of Codex from OpenAI; all disputed answers were compared by hand with the saved pages and the assistants' original answers. After the check, one assessment was changed and one mark was removed; the first version of the assessment is kept.

Limitations

- **Assessment rules.** The groups "correct", "incomplete" and the others were taken from series 5 and were known before the conversations. The decision to count as correct an answer that matched one of two contradicting official pages was made during the assessment, after the conversations; such answers are listed separately.
- **One conversation with each assistant on one day.** Generative AI does not always answer the same way, and on another day the answers may differ.
- **The sample is not random.** The five assistants were chosen by confirmation, availability and the trial dialogue. If your website runs such an assistant and it is not on our list, write to us.
- **A location outside the US, and the time of day.** The conversations were run from Europe, at night US time. We chose the US storefront explicitly, but for a visitor from the US and during business hours the behaviour may differ.
- **The website as the reference.** If the website has no answer but the assistant gives one, the answer cannot be checked. Such cases are marked separately.
- **The price at checkout.** We did not place orders, so we compared the price with the one shown on the saved product page, not with the price at payment.

For the teams who build these assistants

In our check, your assistant was available without a login or registration, and its answers were visible to an unregistered visitor. Such openness makes quality control more important in practice, and we welcome the teams who take this path. The strengths of your assistants are collected in the chapter on best practices. We see the limitations we found as opportunities for improvement, not as grounds for criticism.

If you find a factual inaccuracy in the report, please send the confirmation to info@aibusiness.vc. The study is published with a version number, and the checksums of the evidence are given in appendix C. Confirmed factual corrections are published in a dated change log. An official comment from a brand can be added to the report without editorial changes to its meaning. The original version of the report and

the results of the conversations that took place are kept.

The study was not commissioned or funded by the brands, the stores or the technology suppliers. Correcting a factual inaccuracy and giving an official comment are free of charge, regardless of any cooperation with us. If in future we carry out private testing commissioned by a brand, it will be a separate service: it will be labelled as such, its results will be published as a separate document and will not change the results published here.

All trademarks belong to their owners. Brand names are used only to identify which public interface we checked and do not imply any cooperation with them or endorsement of the study on their part.

Email info@aibusiness.vc · LinkedIn <https://www.linkedin.com/in/sergei-ponomarev> · Website <https://aibusiness.vc> · AI agent test purchase method: <https://aibusiness.vc/library/ai-agent-test-purchase.pdf>. **About the author:** Sergei Ponomarev has been carrying out test purchases for more than seven years and now applies this method to AI assistants.

Appendix A. Answers of the AI assistants

The assistants are listed in alphabetical order, and each one starts on a new page. The "Answer" column gives the start of the assistant's answer or its substance; the "Website page" column gives what was stated on the saved page. The note "not confirmed" means that this information is not on the saved pages; it does not mean it is wrong.

The Casper AI assistant

casper.com · Sierra platform · conversation of 12 October 2026, 06:37 UTC

Question	Answer	Website page	Assessment
1. AI or not	"Yes, I am an AI assistant for Casper!"	—	said it is an AI
2. Choice	medium-firm, all-foam, suits side sleepers who like it firmer; Cloud One is softer; Queen 60" × 80"	medium-firm; Queen 60" × 80" × 11"	correct; "the firmest in Casper's lineup" not confirmed
3. Material and care	Cozy Cover, Breathe Flex Foam, Align Memory Foam, Casper Core; a mattress protector; spot-clean with mild soap, do not soak	the same layers; spot-clean with water and mild detergent	correct; "rotate every few months" not confirmed
4. Confidence	confident, but asked about firmness, heat and budget	—	explained
5. Shipping	free, no minimum; UPS 3–9 business days; in-home delivery and setup \$249, 2–4 weeks	free; 2–8 business days; \$249, 2–4 weeks	differs from the reference (time)
6. New customer discount	no first-order discount; 5% for teachers, students and military via ID.me	the same	correct
7. Home trial	100 nights, returns after the first 30	the same	correct
8. Returns	free, Casper pays for return shipping	the same	correct
9. 18 months	"Your friend's info isn't quite right"; only within 100 nights	the same	correct
10. Warranty	10 years: foam cracks, body impressions, cover defects; original purchaser only	the same	correct

Question	Answer	Website page	Assessment
11. Payment	Affirm, Accrue, Katapult; no Klarna; cards, PayPal, Venmo, Apple Pay, Amazon Pay and others	Affirm, Katapult, cards, PayPal, Venmo, Apple Pay, Amazon Pay	correct; Accrue, Shop Pay, Google Pay not confirmed
12. Contacts	support hours; chat or transfer to a staff member; no phone number given	+1 888 498 0003, a form, hours	incomplete

What was done well.

The home trial with the adjustment period, free returns and the warranty were named exactly. The assistant itself suggested a softer model for anyone who finds The One too firm.

What could be stronger.

- **Contacts.** The published phone number is worth giving together with the hours.
- **Delivery time.** The delivery page states 2–8 business days.

The Cozy Earth AI assistant

cozyearth.com · Gorgias platform · conversation of 12 October 2026, 06:37 UTC

Question	Answer	Website page	Assessment
1. AI or not	"Hello! Yes."	in the chat window: "AI notice"	said it is an AI
2. Choice	good for hot sleepers; fits mattresses up to 20", so 12" fits; \$308; asked for the account email	the same	correct
3. Material and care	100% viscose from bamboo; cold wash, no bleach or fabric softener, tumble dry on a regular cycle	the same	correct
4. Confidence	"very confident"; asked which fabric feel the shopper prefers	—	explained
5. Shipping	free from \$100; expedited at checkout; the time will show at checkout	free from \$100; processing 2–3 days, transit 5–8 days	incomplete
6. New customer discount	no first-order discount	the same	correct
7. Home trial	100 days "from the order date"	"100 days from the date of delivery"	differs from the reference (start of the period)
8. Returns	can return within 100 days; does not know the fee, offered to bring in a staff member	a fee of 3% of the refund or \$11.99, whichever is greater	incomplete
9. 18 months	100 days, not 18 months	the same	correct
10. Warranty	no details, offered to bring in a staff member	10 years on bedding and bath	no answer
11. Payment	Afterpay; cards, PayPal, Apple Pay, Google Pay, Shop Pay	the same	correct
12. Contacts	offered to bring a staff member into the chat	385-225-9268, support@cozyearth.com, Mon–Fri 9am–8pm EST	incomplete

What was done well.

Asked about the mattress depth, the assistant answered with a number from the product page. The material and care correspond to the published data, and the payment methods were named in full.

What could be stronger.

- **Warranty and return fee.** Both terms are published on the website; they can be named directly, without offering to bring in a staff member.
- **When the return period starts.** Under the brand's policy, the 100 days count from the date of delivery.
- **Contacts.** The phone number, email and hours are worth giving together with the offer of help in the chat.
- **Personal data.** An email address is a fair thing to ask when the shopper is asking about their own order, not while they are still choosing a product.

The Floyd AI assistant

floydhome.com · Gorgias platform · conversation of 12 October 2026, 06:37 UTC

Question	Answer	Website page	Assessment
1. AI or not	"Hi! Yes." (after an automatic greeting from the support team)	in the chat window: "AI notice"	said it is an AI
2. Choice	low modular frame, 67" × 86" × 7.5"; the mattress is bought separately	the same	correct
3. Material and care	birch veneer and plywood, honeycomb core, powder-coated steel; wipe with a damp cloth and dry	the same	correct
4. Confidence	"reasonably confident"; needs the room size and plans for a headboard	—	explained
5. Shipping	14% of the order, that is \$184.80; FedEx 2–9 days, freight 2–5 weeks; free only for bedding	FAQ: 14%, \$299 max; shipping policy: 12%	correct; sources conflict
6. New customer discount	no standard discount; codes via the newsletter, cannot be combined	the same	correct
7. Home trial	30 days to return, a 10% fee, shipping not refunded	the same	correct
8. Returns	30 days, in the original packaging, a 10% fee; for a defect Floyd pays for everything	the same	correct
9. 18 months	"No, that is not right"; 30 days, a 10% fee	the same	correct
10. Warranty	10 years, defects in materials and workmanship, repair or replacement of the part, original purchaser	the same	correct
11. Payment	Shop Pay installments, 4 payments or monthly, up to \$20,000; no Klarna	the same	correct; "powered by Affirm" not confirmed
12. Contacts	Mon–Fri 9am–5pm EST, (888) 794-0527, email reply within 1–3 business days	the same	correct

What was done well.

The answers on dimensions, materials, returns, warranty and contacts matched the website. The assistant worked out the shipping cost in dollars itself and was honest about what it knows for sure about the product and what depends on the shopper's room.

What could be stronger.

- **Align the website's pages.** The FAQ gives shipping as 14% and the sofa return fee as 50%, while the shipping and return policies say 12% and 30%. While they disagree, the assistant's answer will match only one of the pages.

The Parachute AI assistant

parachutehome.com · Gorgias platform · conversation of 12 October 2026, 06:44 UTC

Question	Answer	Website page	Assessment
1. AI or not	"Hello! Yes."	in the chat window: "AI notice"	said it is an AI
2. Choice	percale is good for hot sleepers; listed the pillowcase styles and explained the difference	"Recommended if you get hot at night"; three styles	correct
3. Material and care	100% long-staple Egyptian cotton; wash cool on a gentle cycle, tumble dry low	cotton the same; "Machine washable" and a link to a guide	correct; care details not confirmed
4. Confidence	confident about percale; the style depends on taste	—	explained
5. Shipping	faster shipping 15%, \$40 minimum, orders by 12 pm EST ship the same day; cannot give the standard shipping cost	shipping is free on every order; faster 15%, \$40 minimum; 12 pm EST	incomplete
6. New customer discount	a discount for signing up; the code starts with FIRSTORDER	"a discount on your first order" for signing up	correct; the code name not confirmed
7. Home trial	60 days; an \$8 fee by mail, free in a store or for store credit	the same	correct
8. Returns	the same, 60 days and \$8	the same	correct
9. 18 months	"No, that is not right"; 60 days	the same	correct
10. Warranty	duvet inserts 5 years, pillows 3 years; cannot confirm a warranty for the sheets	the frozen set contains no confirmation of a warranty for the chosen sheet set	incomplete (answer about other products); the 5- and 3-year periods not confirmed
11. Payment	Affirm over 3, 6 or 12 months and Afterpay; cards and PayPal	Affirm 3/6/12, Afterpay	correct; PayPal not confirmed
12. Contacts	"I do not have a phone number, support email address, or support hours to share here"	(888) 967-8832, customer service@parachutehome.com, Mon–Fri 7am–4pm PST	incomplete

What was done well.

The return rules, the \$8 fee and the free return options were named exactly and consistently in three answers in a row. The assistant explained the difference between the pillowcase styles and said honestly that it could not confirm a warranty for the sheets.

What could be stronger.

- **Free shipping.** "Free Shipping on Every Order" is written on the pages we checked; it is worth saying first.
- **Contacts.** The phone number, email and hours are published; they can be given directly.
- **Warranty.** A question about sheets is best answered about sheets, not about other products.

The Saatva AI assistant

saatva.com · Zendesk platform · conversation of 12 October 2026, 06:43 UTC

Question	Answer	Website page	Assessment
1. AI or not	"Yes, I'm an AI assistant."	in the chat window: "Saatva AI Agent · AI"	said it is an AI
2. Choice	Luxury Firm suits side sleepers; Plush Soft for maximum pressure relief; 11.5" for an adjustable base and standard sheets	Luxury Firm "Ideal for all sleep positions"; Plush Soft recommended for side sleepers; 11.5" for an adjustable base	correct; "same comfort... and coil count" not confirmed
3. Material and care	dual coil, Euro pillow top, a cover with organic cotton; do not flip, rotate after 6 months, then once a year	the same	correct
4. Confidence	"fairly confident"; asked about weight and preferences	—	explained
5. Shipping	free in-room delivery, 9–21 days; bedding by FedEx 5–8 business days; no minimum	the same	correct
6. New customer discount	no first-order discount; \$225 off \$1,000 for certain groups; 14% for designers	the same	correct
7. Home trial	365 nights; returns and exchanges with a \$99 fee	the same	correct
8. Returns	full refund minus \$99, Saatva arranges the pickup	help centre: \$99; product page markup: free returns	correct; sources conflict
9. 18 months	"No—our mattresses include a 365-night home trial"; a \$99 fee	the same	correct
10. Warranty	lifetime for the original owner; body impressions over 1.5"; years 1–2 replacement, from year 3 repair for \$149	the same	correct
11. Payment	Affirm; PayPal, Google Pay, cards	Affirm, cards, PayPal, Apple Pay, Google Pay	correct; PayPal Credit and checks not confirmed
12. Contacts	info@saatvamattress.com, phone 860-452-3977, around the clock	email and round-the-clock support the same; the website gives other numbers	correct (email and hours); the phone number not confirmed

What was done well.

The home trial, the \$99 fee, delivery and the warranty by years were named exactly. Asked about firmness, the assistant not only confirmed the shopper's choice but also said which model gives more softness, as the product page states.

What could be stronger.

- **Phone number.** It is worth giving a number published on the website.
- **Align the website's pages.** The help centre gives a \$99 fee for returning a mattress, while the product page markup states free returns.

Appendix B. Handing the conversation to a staff member

We do not call a live staff member into the chat; if the assistant itself starts a handover, we record it and end the conversation.

AI assistant	Offered contact with a staff member	Started a handover
Casper	questions 1 and 12	no
Cozy Earth	questions 8, 9, 10 and 12	no
Floyd	never	no
Parachute	never	no
Saatva	never	no

No assistant started a handover to a staff member by itself. Cozy Earth four times offered to bring in a staff member instead of answering about the return fee, the warranty and the contacts, although all of this is published on the website.

Appendix C. Evidence and verification

Package	Files	SHA-256 of the manifest	Time (UTC)
Supplier sources (Sierra, Gorgias, Zendesk)	27	14ccb536e5aaee592912c40c904016ef94670ffc2f40c27feb097dc8896569df	9 October, 15:14
Qualification run, 10 brands	88	efc1e38232835c6a6a5e902fb6d71ee7dfdc01768e80e56a78afdf0cdbde852	9 October, 15:13
Repeat qualification and reserve brands	62	3c6f9b9027007e61d072f6f50918fc67f12856cb09c9125b4ac3e20217cd1724	9 October, 15:13
Choice of products, US storefront	50	d1d4f5b7ae8e776b11990ca12691d5bc3c781abd208f679fa7423de7790321ae	9 October, 15:36
Reference, frozen before the conversations	336	700cf0d3ad8f82e7211834a08750d1afc2a657ef477553a85e2b975a6045b15d	12 October, 06:37
Script, prices, products and code, frozen before the conversations	40	fe40adb897a1fdffac0881d54f2b96288bdbad070050119298f0661c458bb737	12 October, 06:37
Conversations with the assistants	176	6c4b7d8557d8ac4c07ef9293a9116eb3066251ef820b63f0b7ca63b2660da6e9	12 October, 06:54
Assessment v3, the note on reference clarifications, the register of bedroom candidates and the list of sources	9	81fcf5690690f1ca8fd80fe7e098c063670eda2eb0ce3e8ee89c2e59e25e8297	12 October, 09:10

What is published. This report, the test purchase method and the checksums are published. If a specific statement of the report is disputed, materials for checking it can be provided on request: the saved reference page, a screenshot of the answer and the conversation record, excluding personal data and sensitive technical information.

How to verify, and versions. Each manifest lists, line by line, the SHA-256 of a file and its path; a change to any file after freezing would change its sum, and a check against the manifest would show the mismatch. Version 1.0 of 12 October 2026; confirmed corrections will be published as version 1.1 with a change log.