

FIELDWORK NO. 050

Which Humans Can a Company Remove?

Nine companies reduced a human service their customers had noticed.

What separated the ones who got away with it was not the technology, and not the cost.

FIELDWORK NO. 050 • SEPTEMBER 2026

SOURCES: COMPANY POLICY PAGES, SEC FILINGS, INVESTOR CALLS, EXECUTIVE STATEMENTS, CONTEMPORANEOUS REPORTING

Go to the Zappos shipping and returns page and the first line is a boast.

Customers enjoy free shipping, free returns, and 24/7 customer service. The page was last updated on May 26, 2026.

Keep reading and the page gets narrower than the boast.

Merchandise must be returned within 60 days in new condition. Returns scanned by the carrier inside 60 days can go back to the original form of payment. After 60 days, the refund is store credit only, and only up to one year from purchase. Anyone who checked out as a guest gets 60 days and nothing beyond it. Final Sale items cannot be returned at all. Footwear has to come back in the original box without tape or a postal label on it. A return that breaks the rules may earn no refund, or a restocking fee of up to half the item price. If Zappos refunds you and the item never arrives, a warning email goes out at 30 days and a charge lands on your card after 45.

There is also a membership tier now. Zappos VIP members get free expedited shipping.

Almost every guide on the internet still describes Zappos as the company with the 365 day return policy. That is a reasonable summary of what it used to be. The one year window survives, in store credit, for people with accounts.

Now look at what did not change. Shipping is still free everywhere in the United States, including territories and military addresses. Returns are still free the same way. The phone number is still in the footer of every page, 800-927-7671, still described as 24 hours a day, seven days a week, and the FAQ still tells you that you can text CHAT and reach a person that way instead.

Twenty-seven years after launch and seventeen years after Amazon bought the company for \$1.2 billion, Zappos has tightened the return window, restricted guests, changed the currency of late refunds, added a penalty for abuse, and built a paid tier. It has not touched the phone line.

That ordering is the subject of this study.

01 THE OBVIOUS THEORY

The obvious theory, and where it holds

The obvious theory is that human service is expensive, that software gets cheaper every year, and that any company carrying a large human layer is carrying a temporary cost it should be planning to remove.

The theory is not stupid. It is often correct. The clearest case sits in an industry that ran the experiment across every carrier in the world.

SITA trialed the first self-service airline check-in kiosks with Alaska Airlines in 1997. Alaska established web check-in in 1999. Over the following two decades the check-in agent went from the default interaction to the exception, and the industry settled on self-service adoption targets around 80 percent.

The result was not a customer revolt. Passengers who check in through a website, an app, or a kiosk report higher satisfaction than passengers who use a staffed counter. A 2013 study found that human agents scored better on reliability, assurance, and empathy, and found no meaningful difference on responsiveness. Alaska, which moved first, was ranked highest in customer satisfaction in the traditional carrier segment for eight consecutive years.

Airlines removed a customer-facing human interaction that every passenger had experienced, replaced it with software, told everyone they were doing it, and the experience got better.

So the theory works. The question this study is about is why it works there and not elsewhere.

02 CASE

Stitch Fix

Stitch Fix launched in 2011 on a pairing. An algorithm narrowed the inventory and a human stylist chose the five items that went in the box. The company described stylists as the heart of the business, and it sold the stylist to customers directly. At the 2020 peak it employed roughly 8,000 people, about 5,100 of them stylists.

The reduction happened in stages over four years.

In 2020 the company laid off 1,400 stylists in California. In August 2021, shortly after launching Freestyle, its direct-shopping feature, Stitch Fix imposed a minimum 20 hour work week on stylists. Roughly a third of them quit rather than accept it. Some who refused the new schedule were offered \$1,000 to resign in exchange for signing a non-disclosure agreement and a promise not to sue.

Styling quotas rose. Then the quota was replaced by a points system that had stylists producing roughly twice as many fixes an hour as before. The annual report filed in July 2023 showed 5,860 employees and 2,620 stylists, about a quarter of them full time. On March 31, 2024, Stitch Fix eliminated every remaining full-time stylist position. Stylists became part-time contractors.

Active clients fell from about 4.2 million at the peak to roughly 2.3 million by the second quarter of fiscal 2026, a decline of around 45 percent. Fiscal 2024 revenue fell 16 percent to \$1.34 billion with a net loss of \$118.9 million. The company went through three chief executives in the period.

The decline is not attributable to the stylist reduction alone, and anyone who tells you otherwise is skipping past the record. Subscription retail lost momentum across the category over the same years. Stitch Fix also made abrupt changes to its marketing that reduced new customer acquisition independently. The most careful public analysis of the collapse describes a compounding structure rather than a single cause: fewer customers acquired, and the customers who were acquired monetized at lower rates because the service they received had degraded.

One number from the wreckage is worth holding on to. In 2022, while the client base was shrinking, revenue per active client rose 15 percent year over year. The relationship kept working on the customers who still had one. What stopped working was bringing new people into it.

03 CASE

Trunk Club

Trunk Club ran the same idea for men and got to the same place by a different route, which makes the pair more useful than either case alone.

The product was a stylist who sent you a trunk of 10 to 12 items to try at home. Nordstrom bought it in July 2014 for \$350 million, and Erik Nordstrom described the acquisition at the time as moving quickly to evolve with customers. Trunk Club had six physical clubhouses, in Boston, Chicago, Dallas, Los Angeles, Washington and New York, where a consultation with a stylist was free.

In the fall of 2016, Trunk Club started charging \$25 for a home try-on, creditable against a purchase. It had been free. The return window was shortened at the same time. That June the company had announced the closure of its Goose Island distribution center in Chicago, affecting about 250 people.

In the third quarter of that same year Nordstrom took a \$197 million goodwill impairment on Trunk Club, more than half the purchase price. It produced a \$10 million quarterly net loss in a quarter where same-store sales rose. Blake Nordstrom told investors the business had not performed to the expectations held at acquisition. Founder Brian Spaly left weeks later.

The fee did not save it. Six months separated the introduction of the charge and the write-down.

The move that mattered came in 2019, and it is easy to miss because nothing visible was removed. Nordstrom stopped dedicating stylists to Trunk Club and began routing Trunk Club customers to its regular full-line Nordstrom stylists instead. A human was still on the other end. The particular human who had learned your sizes and your objections over three years was not.

In March 2020 all six clubhouses closed and their operations folded into nearby Nordstrom stores, with another \$32 million in charges. On that call Erik Nordstrom said styling was a key differentiator for Nordstrom and described the plan as relocating clubhouse styling into stores to reach more customers while continuing core trunk-based services. Trunk Club closed for good on May 31, 2022.

Nordstrom kept the stylists. It kept calling styling a differentiator, and it has said its styling customers spend seven times more than customers who do not use styling. What it stopped doing was selling styling as a separate product with a separate brand.

The service survived. The company that had been built to deliver it did not.

04 CASE

Klarna

The most recent case is also the fastest, and it is the one where the executive said out loud what the damage was.

In February 2024 Klarna announced, jointly with OpenAI, that its AI assistant had handled 2.3 million conversations in its first month. That was two thirds of Klarna's chat volume, and the company equated it to the work of roughly 700 full-time agents. The reported numbers were strong: resolution in under two minutes against 11 minutes for human agents, a 25 percent drop in repeat inquiries, and about \$40 million in projected annual profit impact. Klarna's headcount went from roughly 5,000 in late 2023 to roughly 3,500 by late 2024, largely through a hiring freeze rather than layoffs. Sebastian Siemiatkowski, the chief executive, said publicly that AI could already do all the jobs humans do.

Those figures were company-reported and published alongside the vendor. They should be read with that in mind. Klarna's own framing was more careful than the headline that circulated, and the company said the 700 figure described hiring it avoided rather than people it removed.

About fifteen months later, in May 2025, Siemiatkowski told Bloomberg the company had gone too far. His account, as reported: they had focused too much on efficiency and cost, the result was lower quality, and that was not sustainable. He added a second thing that is more useful than the first. From a brand and company perspective, he said, it is critical to be clear to your customer that there will always be a human if they want one.

Klarna started recruiting human agents again. What it built was not the old organization. The returning agents work through a flexible remote pool, described

by Siemiatkowski as an Uber-style setup, drawn substantially from students and people with partial availability. The roles that were removed had been full time.

The reversal has stronger independent sourcing than the original claims. Bloomberg, Reuters, and Forbes all covered it.

Note where the chief executive located the harm. Not the product. The brand.

05 CASE

Hilton

The hotel industry ran the same decision at the same time and mostly did it differently, which is why it is in this study.

In July 2021 Hilton announced that in the United States, daily housekeeping at its non-luxury brands would be performed on request rather than by default, with an automatic clean on the fifth day of a stay. Chris Nassetta, the chief executive, described the intent on an investor call in terms nobody had to interpret. The work across every brand was about making them higher-margin businesses with more labor efficiency, particularly in housekeeping and food and beverage, and coming out of the crisis those businesses would be higher margin and require less labor than before. In November of that year he said the company had to communicate with customers and retrain them.

Waldorf Astoria, Conrad, and LXR were exempt from the start. So were the company's Asia Pacific resorts. The withdrawal was written with an edge on it.

The pressure that followed was sustained. The hotel workers' union UNITE HERE delivered a petition with 20,000 signatures and issued a travel alert telling guests that at most Hilton properties the default was no longer a clean room. A housekeeper at Hilton's Drake Hotel in Chicago, 22 years on the job, said guests were stopping her in the hallway to ask why their room had not been cleaned.

Hilton reversed part of it. Daily housekeeping is now standard again across all Hilton luxury, full-service, and lifestyle brands, plus Embassy Suites, with every-other-day service at the remaining brands. Marriott settled into a comparable structure, with full daily service at luxury and resort properties, a daily tidy at the upper-upscale brands, and every-other-day service at select-service brands.

Neither chain automated housekeeping. Neither chain concluded that housekeeping did not matter. They concluded that it mattered at some room rates and not at others, and they wrote the policy along that line. Having written it that way, restoration was available to them, and the restoration ran downward from the top of the price list.

06 THE DISTINCTION

The distinction the cases produce

Set the airline case beside the Stitch Fix case and the difference is not the technology. Both replaced a human interaction with software. Both told customers. One raised satisfaction and one preceded the loss of nearly half a customer base.

The difference is what the human had been doing in the customer's decision.

Nobody has ever chosen an airline because of its check-in agents. The agent was something you got through on your way to a boarding pass. Remove the agent and you have removed a wait. The value the customer wanted was speed and control, and the software delivered more of both than the person did.

People did choose Stitch Fix because a stylist picked the clothes. That was the product description and the marketing claim. Remove the stylist and the box arrives looking similar and containing something the customer did not buy.

There is a third case, and it comes from a company that has nothing to do with either.

In January 2024 Duolingo cut roughly 10 percent of its contract workforce, mostly translators and writers who produced the language exercises, and moved that work onto generative models. The first quarter of 2025 was the strongest in company history by daily active user additions, with more than 10 million paid subscribers and 38 percent year over year revenue growth.

On April 28, 2025, chief executive Luis von Ahn posted an internal memo to LinkedIn announcing an AI-first strategy. The company would gradually stop using contractors for work AI could handle, and headcount would be granted only where a team could not automate further. The memo also said the company would rather move with urgency and take occasional small hits on quality than move slowly and miss the moment.

The backlash was severe enough that Duolingo, a company whose social presence was one of its genuine assets, deleted it. Sentiment inverted across platforms within two weeks. Von Ahn backtracked publicly on May 24, less than a month after posting, saying he did not see AI as replacing what employees do. A year later, on a May 2026 podcast, he said the company had backtracked on the internal rule evaluating every employee on AI usage, and pointed out that Duolingo had never laid off a full-time employee and that headcount grew in the year of the memo.

The contractors Duolingo removed were not a queue and were not a reason. No user had ever encountered them. Nobody installed the app in order to reach a curriculum writer. Their work reached customers; they did not.

So three categories, and they behave differently.

01 A reason is a person the customer chose the company in order to get. The stylist at Stitch Fix and at Trunk Club. The line to a human at Klarna. The Zappos phone number.	02 A queue is a person the customer went through on the way to something else. The check-in agent. The teller for a routine deposit.	03 An invisible production layer is a person whose work the customer consumes without knowing the person is there. Duolingo's contractors.
--	--	--

Automating a queue removes a wait, and customers notice in the direction the company wanted. Automating an invisible layer changes nothing the customer can see, which is why the Duolingo damage came from the announcement rather than from the product, and why the reversal was rhetorical rather than operational. Automating a reason removes the thing that was bought.

What the survivors actually did

Across these cases, and across a search run specifically to find an exception, no company was found that took a human service customers had chosen it for, replaced it with software, and kept the advantage. Two tried. Stitch Fix lost roughly 45 percent of its client base against a background of other causes. Klarna reversed inside fifteen months and its chief executive named the damage as brand.

That is a null result from a bounded search. It is not a law. Nine companies is a small number, one deliberate attempt was made to break the finding, and a case that meets the conditions may exist and may simply not have surfaced.

What the companies that kept their advantage did instead falls into a short list, and none of the entries involve removing the person.

SEGMENT .

Keep the service where the price justifies it and withdraw it where it does not. Hilton exempted its luxury brands on day one and restored downward from there. Zappos built a VIP tier and separated account holders from guests.

METER .

Keep the service and charge for it. Trunk Club's \$25 fee is the cautionary version, since it arrived six months before the write-down and did not change the outcome. HubSpot is the version that worked, because it did it before there was a problem: free and Starter tiers carry no human onboarding and no fee, Professional and Enterprise require paid onboarding, and the separation was designed in rather than retrofitted.

EXTERNALIZE .

Keep the service and move who employs the person. Guidewire has spent roughly two decades on partner-led implementation through certified systems integrators while keeping accountability for the outcome, and its own services revenue still grew 21 percent in the most recent fiscal year on record. HubSpot lets a certified partner take the onboarding work and waive HubSpot's fee, so the customer still pays and Klarna's second act is the same mechanism arriving late, after automation failed.

ABSORB .

Keep the service and move where it lives. Nordstrom kept the stylists and stopped selling styling as a separate product.

TIGHTEN ELIGIBILITY .

Keep the promise and narrow who qualifies under it. This is the quietest device on the list and Zappos is the clearest user of it. The company that would once take back nearly anything now enforces condition standards, flags accounts with unusual return rates, charges up to 50 percent restocking on non-compliant returns, and cuts guests out of the long window entirely. The headline never changed.

Zappos is worth one more sentence, because it is the case most likely to be misread. It did not refuse to cut. It cut hard, and it cut everywhere except the phone line. The economics of that choice are not observable from outside. Zappos has not reported separately since the Amazon acquisition closed in 2009, so there is no public way to know what the service layer costs, what it returns, or what Amazon carries. What can be established is that the promise survived and the subsidy around it did not.

The account of why Zappos built it this way is consistently reported across more than a decade of coverage: Tony Hsieh described taking most of the money that would have gone to paid advertising and putting it into service instead, on the theory that customers would do the marketing. The operating facts line up with that account. No scripts. No upselling. No average handle time target and no calls-per-day quota. The metric was the share of an agent's time spent with customers, and the target was 80 percent. Agents were trained to send callers to competitors' sites when Zappos was out of stock. The call center was never outsourced.

If the service was funded as marketing rather than carried as support overhead, it would not have appeared on the list of costs to cut when the list got written. That is one company and one explanation, and it is offered as a possibility rather than a finding.

08 THE QUESTION

The question underneath all of this

The question an operator reaches for is whether a given human task can be automated. By 2026 the answer to that is usually yes, and the answer keeps getting more yes, and it turns out to predict almost nothing about what happens next.

The question that predicted the outcomes in these cases is different. What was the person doing in the customer's decision to come here?

If the person was a queue, automation is a gift and the only real risk is executing it badly. If the person was never visible, the immediate customer-experience risk is much lower, and the remaining risk may sit in how the change is communicated, which Duolingo learned at considerable cost while its financials kept working.

If the person was a reason, the question stops being about automation. It becomes a question about which of the five moves is available. Can this be segmented so

that the customers who chose the person keep the person? Can it be metered so that the ones who value it pay for it? Is there a partner who could carry it and be paid better than you were charging? Is there a larger business it could live inside? Can eligibility be narrowed without the promise changing?

Most companies do not ask which category they are in, because the human layer usually got built for operational reasons and nobody ever wrote down what it was doing commercially. That is the gap. It is answerable, and it is answerable internally, from data most companies already have. Which customers cite the person in win notes. Whether accounts that used the person churn differently. What happens to conversion when the person is not available.

09 NOW

Where this lands now

A large number of companies founded in the last three years are running the exact structure this study is about. They sell software, they market automation, and they staff a human layer at the front of every customer relationship to make the software land. Implementation specialists. Onboarding leads. Solutions consultants. Reviewers who check the model's output before it reaches anyone. In many cases the human layer is provided free and named in the marketing.

Those companies are all approaching the same fork, and most of them will approach it while growing, which is the worst time to think clearly about it.

Take one live example. DualEntry, an AI-native ERP company that raised \$90 million in October 2025, includes CPA-led implementation on every plan at no charge. Its automated migration moves a customer's data in 24 hours. Its published go-live timeline is four to six weeks, and that part is delivered by people. The company's lead investor reported in its own diligence write-up that multiple customers cited migration speed as the core reason they chose DualEntry, and migration speed is the automated half. Meanwhile customers on review sites have spent more than a year asking for in-app walkthroughs and guided setup, which is a request to need the person less.

Read one way, the CPA is a reason and removing them removes the purchase. Read the other way, the CPA is a queue that customers would happily route around if the product let them, and building that route would be the highest-return work available.

The public record does not settle it, and this study is not going to pretend otherwise. The company can settle it from its own data.

Anyone currently paying for a human layer in front of an automated product has the same week of work available, and a narrower window than they think to do it in. The people who can answer the question tend to be the same people the automation is aimed at.

WHAT THIS STUDY CANNOT ESTABLISH

The conclusion above is a null result from a bounded search. Nine companies were examined and one deliberate attempt was made to find a counterexample. A qualifying case may exist.

Three of the outcomes here are confounded and are presented as mechanisms rather than as measurements. Stitch Fix's client decline ran alongside a category-wide contraction and independent marketing changes, and should not be attributed to the stylist reduction alone. Peloton's move to third-party delivery, which sits in the underlying research and not in this article, coincided with the collapse of pandemic home-fitness demand. Duolingo withdrew daily active user guidance, which removes the cleanest series for reading its outcome.

Klarna's original efficiency figures were company-reported and published with the vendor. The reversal carries stronger independent evidence than the claims it reversed.

Zappos has not reported separately since 2009. Nothing here establishes that the economics of its service layer work. It establishes that the promise survived while the subsidy around it was narrowed.

None of these cases were run as controlled comparisons, and none of the differences between them support a causal claim on their own. What repeats across them is a question that sorted the outcomes, and a question is a smaller thing than a cause.

Sources: Zappos.com shipping and returns policy, page-dated May 26, 2026, retrieved September 4, 2026. Zappos customer service FAQ and site footer, same date. Stitch Fix fiscal 2023 annual report and fiscal 2024 results; contemporaneous reporting in Modern Retail, Retail Dive, Business of Business, and WWD. Nordstrom investor calls of November 2016 and March 2020; Retail Dive, WWD, Fox Business. Klarna and OpenAI joint announcement of February 2024; Klarna Q1 2024 release; Bloomberg, Reuters, and Forbes, May 2025. Hilton investor calls of 2021; The Points Guy, Forbes, CBS News; UNITE HERE press materials. Duolingo Form 8-K for the first quarter of fiscal 2025; Luis von Ahn's April 2025 memo as posted by him; contemporaneous coverage. SITA and Alaska Airlines self-service history; passenger satisfaction research on self-service check-in. Guidewire and HubSpot findings carried forward from prior fieldwork on enterprise deployment.