Will AI Replace Human Sales Reps? An Evidence Review of AI vs. Human Agents in B2B Phone Sales

Augment, not replace, yet: what the peer-reviewed evidence reveals about AI sales agents, telesales automation, and conversation intelligence

Aleksei Prishchepo

2026-09-12

Contents

Summary 1

How this review was done 2

About the author 2

Finding 1: AI is already capable of replacing human reps 3

Finding 2: But replacement is not yet viable 3

Finding 3: Augmentation is where AI creates measurable value 3

Finding 4: Why the acceptance gap exists, and why it may narrow 4

Discussion and implications 4

Limits of this review 4

Verdict 5

Studies reviewed 6

References 7

This is a synthesis of published research, not primary market research.

Summary

Can AI replace human sales reps in high-volume, low-ticket B2B phone sales? The evidence available today points to two different answers. AI is already **capable**: in a controlled field experiment, an undisclosed AI agent making outbound sales calls converted customers at the same rate as proficient human rep. But it is not yet **viable** as a replacement: the moment the agent's machine identity is disclosed up front, purchases collapse by more than 79.7%. Meanwhile, every deployment study that measured outcomes found the largest and most reliable gains not when AI replaced the rep, but when it **supported and coached** the rep. The most viable and adoptable use of AI in these conversations today is augmentation, not replacement. That is a statement about the present, not a prediction that it holds forever.

How this review was done

This review was designed to challenge its own conclusion, not to confirm it. The method:

- **Falsifiable question.** We stated the founder-favourable proposition as the null hypothesis ("AI is not rapidly replacing human reps in this segment right now") and and set out to *reject* it rather than gather evidence in its support. Kill criteria were fixed in advance: evidence that AI is both capable and accepted without penalty, that resistance is declining rapidly, or that a real-world deployment had replaced human reps at scale with equal or better results.
- **Source pipeline.** Candidate sources were gathered, tabulated, and then audited for relevance to the specific segment (B2B, low-ticket, high-volume, phone), for quality (with peer-reviewed and independent sources preferred), and for conflicts of interest. Any source with a commercial stake in the answer was excluded, **regardless of which side it supported**: AI vendors who benefit from "AI wins" and human-agent or sales-tech vendors who benefit from "humans win" were both excluded. Vendor blogs, marketing case studies, and non-primary write-ups were removed.
- **Reported figures.** Numbers are quoted as the primary source states them. Where a figure could not be located in the source, it is not reported.

The result is a small pool of studies: the honest conclusion is that rigorous evidence specific to this segment is scarce.

About the author

Aleksei Prishchepo[1] is an independent BI and analytics consultant based in Berlin. He came to analytics after more than two decades running businesses, where he worked with sales, operations, finance, and the decisions that connected them.

His approach to analytics starts with the business question rather than the dashboard: define what would count as evidence, make the assumptions explicit, and look for reasons an analysis could be wrong. This report is an example of that approach[2].

He also builds Sales Signals[3], a sales analytics product that operates in the market examined here.

[1] Aleksei Prishchepo on LinkedIn: linkedin.com/in/aleksei-pr
[2] Aleksei's BI and analytics work: frequentist.org
[3] Sales Signals: salessignals.app

Finding 1: AI is already capable of replacing human reps

In a randomized field experiment involving 6,255 customers, an **undisclosed AI agent converted 23.7% of customers**. This was **statistically indistinguishable from the 25.1% conversion rate of proficient human reps** ($p > 0.10$) and roughly **four times the rate of inexperienced reps** (4.9%) (Luo *et al.*, 2019).

In terms of conversational performance alone, the machine already matches a proficient human at near-zero marginal cost.

Finding 2: But replacement is not yet viable

The same experiment also revealed the barrier. When the agent disclosed that **it was a machine before the conversation, purchase rates fell by more than 79.7%** (to 4.8%), and **56.3%** of customers hung up (Luo *et al.*, 2019). The barrier is not capability; it is acceptance. And because disclosure is increasingly an ethical and regulatory requirement, that capability advantage is lost at the point of deployment.

How the agent sounds and behaves also matters. The acoustic characteristics of human and AI sales agents' voices measurably affect customer purchases (Tong *et al.*, 2023), and adding human-like vocal fillers ("um," "hmm") to a commercial voice agent *reduced* customers' verbal responses (from 6.5% to 3.5%, $p = .005$) and purchase interest (from 2.1% to 0.1%, $p = .006$) by triggering suspicion of manipulation (Liu, Liu and Zhu, 2024) (field experiment, ~4,939 robocalls / ~1,683 connected).

Finding 3: Augmentation is where AI creates measurable value

Across the deployment studies, the pattern is consistent: AI creates the most value as a tool for the rep, not as a replacement for the rep.

- **Sequential funnels.** In a field experiment involving ~78,169 prospects in SME lending, AI handled outbound screening while human reps handled the closing stage: the AI-initiated funnel **increased the number of completed applications**, with the gains appearing specifically **during the human follow-up phase** (Li *et al.*, 2026).
- **AI coaching.** AI coaching lifted **middle-ranked** agents most (about **+8.02%**) versus smaller gains for bottom (**+4.06%**) and top (**+1.00%**) agents relative to human managers (an inverted-U); a restricted-AI variant raised bottom agents by **+4.64%**, roughly **+50% relative**. AI-plus-human coaching produced substantially larger improvements for lower-performing agents than AI alone, whereas top performers saw minimal gains from the combined approach and actually declined under AI-only coaching (Luo *et al.*, 2021).

- **AI-augmented creativity.** AI-assisted lead generation increased reps' success in handling untrained customer questions by about **2.33×** and roughly **doubled purchases (4.5% vs. 2.8%, $p < 0.05$)**; AI alone reached **2.1%**, matching bottom reps (**2.0%**) but trailing top reps (**3.7%**). The effect was **more pronounced for higher-skilled reps** (Jia *et al.*, 2024) (field experiment, ~3,144 customers / 40 reps).

Two qualifications are important here. First, the studies disagree on *who* benefits most (the coaching study points to middle performers, the creativity study to higher-skilled reps); these are different interventions, so both can be true, but neither supports a blanket "AI rescues weak reps." Second, augmentation is not frictionless: the same literature documents skill-biased effects and organisational strain when AI concentrates work on senior reps.

Finding 4: Why the acceptance gap exists, and why it may narrow

Research on the underlying mechanisms helps explain the acceptance gap. People show **algorithm aversion** in judgement-laden, relational tasks, but that aversion is not fixed: it eases with prior exposure and, notably, when people can even slightly **modify** the algorithm's output (Dietvorst, Simmons and Massey, 2018; Jussupow, Benbasat and Heinzl, 2024). This suggests that the barrier may not be permanent. If customer familiarity with AI continues to rise, or disclosure is shifted to later in the call, the acceptance penalty that currently protects the human rep could erode.

Discussion and implications

The most defensible interpretation: **today, in B2B phone sales, the AI that gets adopted and creates measurable value is AI that augments human reps, not AI that replaces them.** Replacement is technically possible, but currently constrained by an acceptance penalty that is real and reinforced by disclosure requirements.

The strategic implication is not "humans are safe forever." It is the opposite: the capability already exists, so the durable position is to build on the augmentation use case that works today while accumulating the conversational data that will matter regardless of how customer acceptance evolves.

Limits of this review

- The empirical studies pre-date current generative-LLM voice agents; they test earlier, more scripted systems. Modern agents may face a different acceptance curve.
- None of the studies examines the exact target context (wholesale/distributor reorder desks); the closest analogues are financial-services and telemarketing outbound sales.

- Evaluation horizons are short (single calls to about one month), so long-run habituation is not directly measured.
- This is a synthesis of a small number of published studies, not primary research. It is strongest as a map of what is currently known and where the evidence is thin.

Verdict

The pre-committed test was whether the evidence showed that AI was both capable of replacement and accepted by customers, that resistance to AI sales agents was declining rapidly, or that AI had replaced human reps at scale with equal or better results. **None of these criteria was met.** We therefore **do not reject** the hypothesis that human reps are not being rapidly replaced in this segment today. The evidence base is small and partly dated, so treat this as the most defensible reading of what is currently known, not a settled answer. It is still enough to act on: augmentation is the use that works now, and it builds the conversational data that will matter however customer acceptance shifts.

Studies reviewed

Study	Role in review
Xueming Luo and Siliang Tong and Zheng Fang and Zhe Qu (2019). "Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases."	Core evidence: capability parity and disclosure penalty
Xiaolin Li and Qing Li and Danny Wang and Flora Gu (2026). "AI-Human Collaboration in Sequential Sales Funnels: Evidence from A Large-Scale Field Experiment."	Augmentation: AI screening and human closing
Siliang (Jack) Tong and Xueming Luo and Erya Ouyang and Jaakko Aspara and Zheng Fang (2023). "How the Voice of Human vs. AI Sales Agents Impacts Customer Purchases."	Voice and acoustic cues in AI vs. human sales calls
Guilin Liu and Maggie Wenjing Liu and Qichao Zhu (2024). "Hmm, the effect of AI conversational fillers on consumer purchase intentions."	Humanising tactics: vocal fillers backfire
Xueming Luo and Marco Shaojun Qin and Zheng Fang and Zhe Qu (2021). "Artificial Intelligence Coaches for Sales Agents: Caveats and Solutions."	AI coaching of reps: effects vary by skill level
Nan Jia and Xueming Luo and Zheng Fang and Chengcheng Liao (2024). "When and How Artificial Intelligence Augments Employee Creativity."	AI lead generation augments reps, with skill-dependent effects
Ekaterina Jussupow and Izak Benbasat and Armin Heinzl (2024). "An Integrative Perspective on Algorithm Aversion and Appreciation in Decision-Making."	Mechanism: when people accept or reject algorithms
Berkeley J. Dietvorst and Joseph P. Simmons and Cade Massey (2018). "Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them."	Mechanism: aversion erodes with control and exposure

References

Dietvorst, B.J., Simmons, J.P. and Massey, C. (2018) "Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them," *Management Science*, 64(3), pp. 1155–1170. Available at: https://doi.org/10.1287/mnsc.2016.2643.

Jia, N. *et al.* (2024) "When and how artificial intelligence augments employee creativity," *Academy of Management Journal* [Preprint].

Jussupow, E., Benbasat, I. and Heinzl, A. (2024) "An integrative perspective on algorithm aversion and appreciation in decision-making," *MIS Quarterly*, 48(4), pp. 1575–1590. Available at: https://doi.org/10.25300/MISQ/2024/18512.

Li, X. *et al.* (2026) "AI-human collaboration in sequential sales funnels: Evidence from a large-scale field experiment," in *Proceedings of the 48th annual ISMS marketing science conference (ISMS 2026)*. Carcavelos, Portugal. Available at: https://scholars.hkbu.edu.hk/en/publications/ai-human-collaboration-in-sequential-sales-funnels-evidence-from-.

Liu, G., Liu, M.W. and Zhu, Q. (2024) "Hmm, the effect of AI conversational fillers on consumer purchase intentions," *Marketing Letters* [Preprint]. Available at: https://doi.org/10.1007/s11002-024-09760-4.

Luo, X. *et al.* (2019) "Frontiers: Machines vs. Humans: The impact of artificial intelligence chatbot disclosure on customer purchases," *Marketing Science*, 38(6), pp. 937–947. Available at: https://doi.org/10.1287/mksc.2019.1192.

Luo, X. *et al.* (2021) "Artificial intelligence coaches for sales agents: Caveats and solutions," *Journal of Marketing*, 85(2), pp. 14–32. Available at: https://doi.org/10.1177/0022242920956676.

Tong, S.(Jack). *et al.* (2023) "How the voice of human vs. AI sales agents impacts customer purchases," in *Proceedings of the 2023 conference on artificial intelligence, machine learning, and business analytics*. Philadelphia, PA, USA. Available at: https://research.hanken.fi/en/publications/how-the-voice-of-human-vs-ai-sales-agents-impacts-customer-purcha/.