

When should an AI sales agent disclose product limitations?

Student: illustrative example | Degree: Master's | Area: Human-AI interaction / sales agents
Discussion outline, not an official AIT proposal form. Proposed study; no data collected.

Background

AI sales agents should help customers understand product fit, including limitations. Prior work shows that reducing overreliance can involve trade-offs in user experience [1], and language models can favour agreement over truthfulness [2]. These adjacent findings motivate the question; they do not establish a gap in sales-agent research.

Problem

A factually correct recommendation may still leave customers with a poor understanding of exclusions. We propose to test whether the **timing of the same limitation disclosure** affects understanding and choice. A targeted review of sales, disclosure timing, and conversational recommendation research is still needed.

Research question

In simulated insurance-product conversations, does presenting a key exclusion **before** rather than **after** a recommendation improve customers' understanding and their selection of a product that meets stated needs?

Proposed approach

Build a small sales-agent prototype using fictional product descriptions checked against a fixed answer key. Compare two disclosure orders using the **same facts, wording, recommendation, and interface**. Both conditions disclose the exclusion before the customer chooses; only its position changes. This tests disclosure timing, not an adaptive-agent mechanism.

Evaluation

Randomly assign participants to one of the two orders for a fixed scenario. Primary measure: exclusion comprehension (0-6, using prewritten factual questions). Secondary measures: choice that satisfies the scenario's explicit requirements, perceived pressure, and task time. Keep the comprehension quiz after the choice to avoid prompting it. Predefine scoring and analysis; report effect sizes and uncertainty. Check reading level, prior familiarity, and whether participants noticed the disclosure. A simulated task cannot establish real purchase benefit.

Expected contribution

An **anticipated empirical finding** about when truthful limitations should appear in a sales conversation, plus a reusable scenario and evaluation protocol. A null result would also inform design. Novelty and practical significance remain to be assessed; no findings are claimed.

Feasibility and first step

Scope: an illustrative master's project within the lab's sales-agent work. First compare the closest literature and prepare one fictional scenario. After the required ethics review, pilot with 8-12 participants to test clarity and scoring, not effectiveness. Use pilot variability and a justified target effect to plan the main sample. Maintain the paper, code, and decisions in the project repository.

Starting references - adjacent work, not a completed novelty review

[1] Bucinca, Malaya & Gajos (2021). *To Trust or to Think*. Proc. ACM HCI. doi:10.1145/3449287

[2] Sharma et al. (2023; revised 2025). *Towards Understanding Sycophancy in Language Models*. arXiv:2310.13548