REMEMBER THE ALTERNATIVES: CHOICE-CONTEXT MEMORY FOR PERSONALIZED DECISION AGENTS

Anonymous authors
Paper under double-blind review

ABSTRACT

An embodied agent receives the instruction *sit somewhere* in a living room. Several seats may all satisfy the instruction and the supplied feasibility constraints. Once these explicit constraints are satisfied, the remaining ambiguity is personal: which seat does this user prefer? Past choices can help resolve it. Choosing a nearby wide chair over a distant chair may suggest that distance matters in that comparison. Choosing the same chair over a nearby narrow seat may instead provide evidence for a preference over seat width. These observations do not uniquely identify the user's preference, but they carry different information about it. If memory stores only the selected wide chair, the two interactions become indistinguishable. Recent personalized agents increasingly use persistent memory to adapt across interactions. Most work focuses on how stored experiences are organized, updated, and retrieved. We instead ask what information a choice event should preserve before it enters memory. We identify a failure we call *choice-evidence aliasing*: different choice contexts can collapse into the same stored outcome, even when they provide different evidence about user preferences and support different future actions. We formalize the information lost by this compression and show that it can change a later optimal decision. This motivates *choice-context memory*, which preserves the selected target together with the alternatives. Building on this principle, we introduce AFFORDANCEMEM, a personalized decision agent for human–scene interaction. It combines current instruction and feasibility constraints with preference beliefs inferred from past choice context, then decides whether to execute a target, ask for clarification, or defer. Clarification responses are stored as new choice-context memories. We evaluate the framework on HUMANISE–ScanNet and LINGO with controlled preferences and repeated interactions. Matched comparisons and targeted interventions show that preserving choice context improves later personalized decisions and can change whether the agent acts or asks. These effects persist across choice mechanisms, interaction sessions, and scene constructions. Our results support choice context as a decision-relevant memory unit for personalized agents.

1 INTRODUCTION

Consider an embodied agent that receives the instruction *sit somewhere* in a living room. Several seats may satisfy the instruction and all supplied feasibility constraints. Once a target is specified, a scene-conditioned motion executor can realize the interaction. The unresolved decision is which valid seat the agent should choose for this user. Different users may prefer different seats in the same scene. A personalized agent should therefore learn from earlier choices and avoid asking again when those choices already provide enough evidence.

Past choices can help, but only if memory preserves why they were informative. Suppose chair A is nearby and relatively wide. Choosing A over an equally wide but distant chair may provide evidence that distance mattered in that comparison. Choosing the same chair over an equally near but narrower seat may instead provide evidence about width. Neither observation identifies the user's preference on its own, but they do not carry the same evidence. A memory that stores only chair A merges the

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

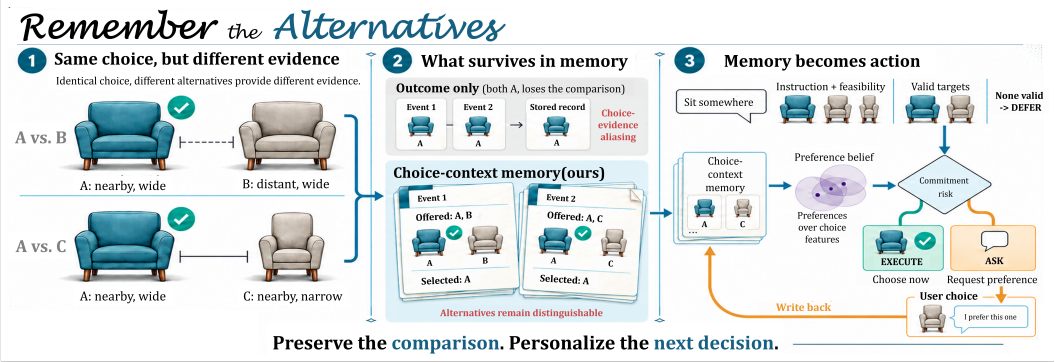

Figure 1: **Remember the alternatives.** Two interactions can share the same selected target but provide different evidence when their alternatives differ. AFFORDANCEMEM stores each choice with its offered alternatives, uses this choice context to infer user preferences, and turns the resulting evidence into a later EXECUTE, ASK, or DEFER decision.

two experiences. It remembers the outcome while discarding the comparison needed to interpret it. When the agent later chooses between a nearby narrow seat and a distant wide one, this missing distinction can lead to a different decision.

Persistent memory is increasingly central to personalized agents. Recent work studies how conversational histories should be divided and retrieved (Pan et al., 2025), how memories can be linked and updated as experience accumulates (Xu et al., 2025), and which abilities long-term memory systems should support (Wu et al., 2025). Interactive personalization also studies when an agent should actively discover missing user preferences (Li et al., 2026). These directions have improved how interaction histories are constructed, organized, retrieved, and extended. We focus on a more specific question for choice-based feedback: when a user choice is written to memory, what information should the record preserve? We do not ask whether an agent should remember the interaction. We ask what must survive when that interaction becomes memory.

The missing piece is the comparison behind the choice. When memory retains only the selected outcome, choices made against different alternatives can become the same record. The outcome is preserved, but the evidence carried by the comparison is lost. We call this failure *choice-evidence aliasing*. To preserve this evidence, we introduce *choice-context memory*, which stores the selected target together with the alternatives that were available at the time. The goal is not to infer a user's preference from a single choice. Instead, the memory preserves the context that can help interpret that choice together with later interactions. We formalize this distinction and show that histories with the same selected outcome can support different future actions when their original alternatives are retained.

The next question is how the agent should use this memory. Past choices may provide enough evidence to favor one valid target, in which case the agent can act directly. In other cases, several targets may still be consistent with what the agent has learned about the user, and asking for clarification is more appropriate. Choice context therefore supports two related decisions: which valid target best fits the user, and whether the available evidence is sufficient to act without asking again.

Building on this idea, we introduce AFFORDANCEMEM, a personalized decision agent for human–scene interaction. For each new scene and instruction, the agent first applies instruction and feasibility constraints to determine which targets are valid. It then uses past choice-context memories to form a belief over the user's preferences and select among the remaining targets. The agent executes a target when the evidence is sufficient, asks for clarification when more information is needed, and defers when no valid target exists. A clarification response is stored as a new choice-context memory and can inform later interactions. We instantiate the preference belief with both a finite hypothesis bank and a conditional normalizing flow. Both implementations share the same memory and decision interface. Candidate proposal and motion execution remain supplied components, as illustrated in Figure 1.

Our contributions are:

- **A choice-memory problem.** We identify *choice-evidence aliasing*, where different choice contexts become indistinguishable when memory retains only the selected outcome. We formalize how this lost comparison can change a later decision.
- **AFFORDANCEMEM: a memory-to-action personalized decision agent.** We introduce *choice-context memory* and AFFORDANCEMEM, which connects retained choice evidence to a preference belief and to selective EXECUTE, ASK, and DEFER decisions under current instruction and feasibility constraints.
- **A controlled evaluation of memory-to-action effects.** We isolate the contribution of historical alternatives with parameter-matched models and targeted interventions, then test whether the same principle persists across repeated interactions, different choice mechanisms, and different scene and target constructions.

2 RELATED WORK

Agent memory and personalization. SeCom studies how conversational histories should be segmented and retrieved (Pan et al., 2025), while A-MEM organizes memories as linked records that evolve with new experience (Xu et al., 2025). AdaMem moves further upstream and studies what information should be written into persistent memory in the first place (Chen et al., 2026). LongMemEval evaluates memory across conversational sessions (Wu et al., 2025), and AMemGym extends this direction to interactive, on-policy evaluation of memory-driven personalization (Cheng et al., 2026). Personalized agents also learn from ongoing user feedback: PrefEval studies preference following (Zhao et al., 2025), PrefDisco (Li et al., 2026) and GATE (Li et al., 2025) study proactive elicitation, and ProPerSim studies continual adaptation through user–assistant interaction (Kim et al., 2026). These works improve what agents remember, how memories are managed, and how feedback is acquired over time. We focus on a different level of the problem: what comparison evidence a choice record should preserve so that it remains useful for a later decision.

Learning from choices and alternatives. Random-utility models condition a selection on its offered set (Luce, 1959; McFadden, 1974; Negahban et al., 2018). Context-dependent models account for effects of the alternatives themselves (Seshadri et al., 2019; Zhang et al., 2025). Robot reward inference also reasons about demonstrations a user could have provided (Jonnavittula & Losey, 2021) and the consequences of a misspecified choice set (Freedman et al., 2020). Our setting takes the logged offered set as given and asks whether its unchosen members remain available after the choice becomes memory. Blackwell's ordering supplies the information comparison (Blackwell, 1953); matched learners and interventions test the consequences for later decisions.

Selective action and preference elicitation. Selective prediction and learning to defer connect uncertainty with the decision to withhold an output (El-Yaniv & Wiener, 2010; Geifman & El-Yaniv, 2019; Verma & Nalisnick, 2022). KnowNo asks for assistance under planning uncertainty (Ren et al., 2023). Regret-based preference learning focuses questions on useful behavior (Wilde et al., 2020). We use the same decision-centered view: uncertainty matters when it changes the current valid action or makes a wrong commitment costly.

Human–scene interaction. HUMANISE aligns language, motion, and scanned scenes (Wang et al., 2022). SceneDiffuser studies scene-conditioned generation and planning (Huang et al., 2023); COINS controls semantic contacts (Zhao et al., 2022). TRUMANS (Jiang et al., 2024b) and LINGO (Jiang et al., 2024a) expand dynamic and multi-stage interaction. Other models connect text, scenes, and motion (Cen et al., 2024; Wang et al., 2025). AFFORDANCEMEM supplies a personalized target to an executor, which lets us study the preceding decision independently of motion synthesis.

3 AFFORDANCEMEM: FROM CHOICE EVIDENCE TO ACTION

The main design choice in AffordanceMem is what the agent preserves from each past choice: not only the selected item, but also the alternatives that made the choice informative. We organize personalized decisions around a simple principle: preserve the comparison behind a choice, then use that evidence to guide the next action. AFFORDANCEMEM turns this principle into a decision

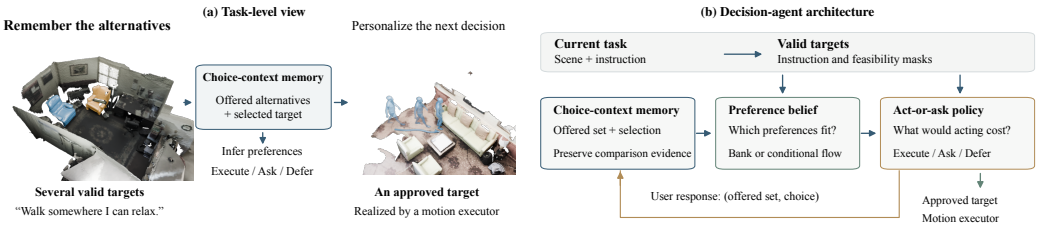

Figure 2: **From choice context to selective action.** AFFORDANCEMEM stores a selected target together with its offered alternatives, uses this evidence to infer a belief over user preferences, and decides whether to EXECUTE, ASK, or DEFER under the current constraints. Explicit user responses update memory, while approved targets are passed to the supplied motion executor.

loop with three parts: choice-context memory, a preference belief, and an act-or-ask policy. Figure 2 summarizes the full pipeline.

3.1 DECISION SETTING

At interaction t, the agent receives a scene, an instruction, and a finite candidate set C_t. A candidate target is denoted by $c \in C_t$, with observable features $\phi(c)$. Binary instruction and feasibility masks, $I_t(c)$ and $F_t(c)$, determine which candidates are currently valid:

$$V_t = \{c \in C_t : I_t(c) = 1 \wedge F_t(c) = 1\}. \tag{1}$$

Thus, C_t describes the full candidate context, while $V_t \subseteq C_t$ contains only targets that the agent is allowed to execute. User preference is considered only within V_t and never overrides these hard constraints. Candidate proposal, masks, and motion execution are supplied interfaces.

The agent may execute a valid target, ask the user for clarification, or defer when no valid target exists. We consider one clarification before the current action. Its response is stored and can inform later interactions. For readability, we omit the interaction index t within a single decision and write the current candidate context as C and its valid target set as V.

3.2 CHOICE-CONTEXT MEMORY: PRESERVE THE COMPARISON

A historical choice event contains an offered set S and the user's selected item $y \in S$. We store it as $m = (S, y)$, together with the candidate features, and write the history as $H = ((S_k, y_k))_{k=1}^{n}$. Here, S_k is the set shown in the k-th past interaction and y_k is the item selected from it. A chosen-only memory keeps y_k and its features but discards the unchosen alternatives in S_k.

Choice-evidence aliasing. Two events can share the same selected item while providing different evidence about the user. If memory keeps only the selected item, these events become the same record. We call this *choice-evidence aliasing*. The missing information is the comparison against the alternatives, and that comparison can matter when the histories support different later actions.

The following proposition states the basic information guarantee. Let h denote the user's latent preference, and let $L(a, h)$ be the loss of taking action a under preference h. For any observed record T, define $\mathcal{B}(T)$ as the lowest expected downstream loss achievable by a decision rule that observes T.

Proposition 1 (Decision value of choice context)**.** *Let $X = (S, Y)$ be a complete historical choice event and let $G = g(X)$ be its chosen-only record. Under any joint distribution over the latent preference h and the observed event X, and for any bounded downstream loss,*

$$\mathcal{B}(X) \leq \mathcal{B}(G). \tag{2}$$

The inequality can be strict: two events with the same selected item can support different optimal future actions.

The result is intuitive. A decision rule with access to X can always discard S and reproduce any rule that uses only G. The reverse is not always possible because the missing alternatives cannot

generally be recovered. Appendix A provides the proof and a finite example. The design implication is simple: a memory record should preserve the comparison evidence, either as the full event or through a representation that retains the same decision-relevant information.

3.3 PREFERENCE BELIEF: INTERPRET THE COMPARISON

A single choice rarely identifies one unique preference. We therefore maintain a belief over possible preference hypotheses h. For a candidate c in context C, $v_h(c; C)$ denotes the utility assigned to c under hypothesis h. The utility may depend on the other candidates in the same context. Each hypothesis also has a positive temperature τ_h, which controls how strongly the user is expected to favor higher-utility items. We use S for a historical offered set and C for the current candidate context; the same utility form can be evaluated in either context.

For a historical choice (S, y), we use the standard choice likelihood (Luce, 1959; McFadden, 1974)

$$\begin{aligned} p(y \mid S, h) &= \frac{\exp(v_h(y; S)/\tau_h)}{\sum_{c \in S} \exp(v_h(c; S)/\tau_h)}, \\ q(h \mid H) &\propto p_0(h) \prod_{k=1}^{n} p(y_k \mid S_k, h), \end{aligned} \tag{3}$$

where $p_0(h)$ is the prior over preference hypotheses and $q(h \mid H)$ is the belief after observing history H. The alternatives in S enter the likelihood directly. They therefore affect how the same selected item updates the preference belief.

Two belief implementations. AFFORDANCEMEM-BANK represents h with a finite set of hypotheses. Each hypothesis contains a preference-weight vector z and a choice temperature τ. Its utility is $v_h(c; C) = z^\top \widetilde{\phi}_C(c)$, where $\widetilde{\phi}_C(c)$ is the feature vector of candidate c standardized within the current candidate context C. Equation 3 updates the hypothesis weights, and expectations are computed by weighted summation.

AFFORDANCEMEM-NF replaces the finite bank with a learned continuous belief $q_\theta(z \mid H)$ over preference vectors z. An order-invariant encoder summarizes the selected items, the unchosen alternatives, and their feature differences, following Deep Sets (Zaheer et al., 2017). A conditional normalizing flow (Dinh et al., 2017) maps a Gaussian base distribution to preference samples, and a contextual decoder scores the current candidates. Training predicts held-out choices from earlier events; Appendix B.3 gives the objective and implementation details.

The two backends differ in how they estimate the preference belief, but they use the same choice-context memory and the same downstream decision rule. This separation is important to our study: in the matched flow comparison, the parameter count, decoder, and training objective are fixed, while only the historical evidence is changed.

3.4 ACT-OR-ASK POLICY: USE THE EVIDENCE FOR THE CURRENT DECISION

The preference belief is useful only through the decision it supports. For each valid target $i \in V$, we measure the expected utility lost by choosing i instead of the best valid target under each possible preference:

$$\begin{aligned} R_q(i; C, V) &= \mathbb{E}_{h \sim q}\left[\max_{j \in V} v_h(j; C) - v_h(i; C)\right], \\ i^* \in \arg\min_{i \in V} R_q(i; C, V), &\qquad \rho(q; C, V) = R_q(i^*; C, V). \end{aligned} \tag{4}$$

Here, $R_q(i; C, V)$ is the expected regret of committing to target i, i^* is the valid target with the lowest expected regret, and ρ is the remaining regret after choosing this best current target. A small ρ means that the preference hypotheses largely agree on what the agent should do.

We compare ρ with a threshold λ, which controls when the agent acts and when it asks:

$$a(H, C, V) = \begin{cases} \textsc{Defer}, & V = \varnothing, \\ \textsc{Execute}(i), & V = \{i\}, \\ \textsc{Execute}(i^*), & |V| \geq 2,\ \rho(q; C, V) \leq \lambda, \\ \textsc{Ask}, & |V| \geq 2,\ \rho(q; C, V) > \lambda. \end{cases} \tag{5}$$

The first two cases are settled by hard constraints: there is either no valid target or only one. Preference memory is used only when several valid targets remain. In that case, the agent executes when the expected cost of commitment is sufficiently low and asks otherwise. When clarification has cost c_{ask} measured on the same decision-loss scale, the one-step rule corresponds to $\lambda = c_{\text{ask}}$ under a target-resolving response. In our experiments, λ is fixed on validation data.

An explicit clarification response is stored in the same form as any other choice event, (S, y), before the decision is evaluated again. Thus, the same representation connects past experience, newly acquired feedback, and future personalized decisions. Together, these components turn preserved comparison evidence into a selective personalized action, which is the central memory-to-decision interface studied in this work.

4 Experiments

We evaluate four questions. Does choice context improve later target selection when model capacity is fixed? Does erasing unchosen alternatives change the agent's later decision? Can the resulting preference evidence support selective action? Finally, does the same principle persist across interaction sessions, choice mechanisms, and scene constructions?

4.1 Experimental setup

Decision episodes. HUMANISE-aligned ScanNet scenes provide candidate objects (Wang et al., 2022; Dai et al., 2017). Each candidate has eight geometric features: root-relative x/y position, planar distance, three box dimensions, height above the floor, and distance to the room center. Each episode contains 6 or 10 historical choice events, a current candidate set, and instruction and feasibility masks.

Controlled preference functions generate historical choices and provide the hidden best valid target and utility regret for evaluation. Models are trained and selected only from observed choices. We split scenes before constructing episodes, yielding 25,756, 7,256, and 5,288 train, validation, and test episodes from 239, 62, and 51 scenes. We also test four shifted preference functions and six non-multinomial-logit (non-MNL) choice mechanisms.

Comparisons and metrics. We compare no memory, chosen-only memory, MAP and Laplace multinomial logit (MNL), and learned Gaussian and normalizing-flow beliefs. Our main representation comparison uses parameter-matched NF models that differ only in whether historical alternatives are visible.

We report hidden-target accuracy and utility regret. For selective action, we report autonomous coverage and error on the autonomous subset. The evidence-erasure study uses a separate 256-particle reference posterior to identify Execute-to-Ask transitions. Confidence intervals use paired scene-level bootstrap samples. Additional details are given in Appendix B.

4.2 Choice context improves later target selection

On 1,322 held-out free-choice sources, the choice-context NF reaches 77.00% hidden-target accuracy, compared with 58.07% for the matched chosen-only NF (Figure 3a). Across six non-MNL mechanisms, the corresponding results are 75.97% and 56.53%. The models use the same current candidates, decoder, objective, and number of parameters; only the historical memory differs.

A choice-context Gaussian model also performs strongly (Appendix C), indicating that the retained comparison evidence is not tied to one belief parameterization.

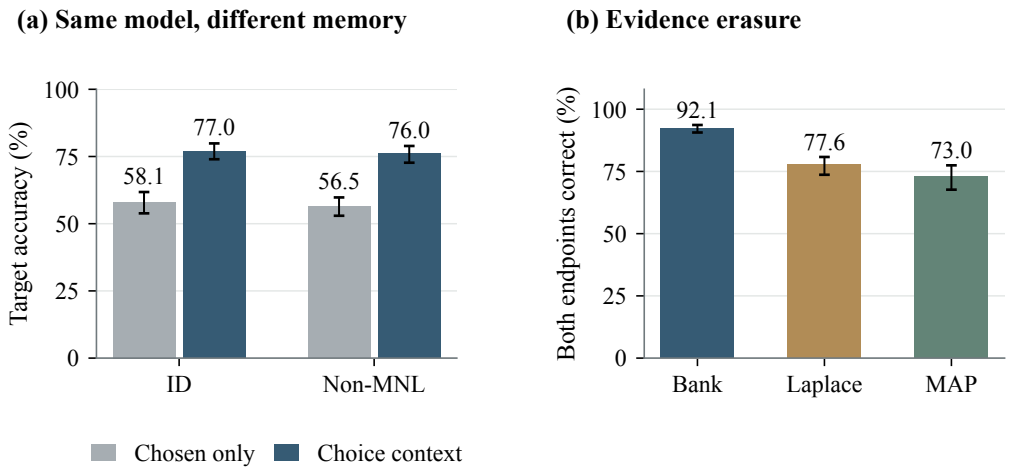

Figure 3: **Historical alternatives carry usable decision evidence.** (a) Parameter-matched NF models differ only in historical memory. (b) Recovery of both endpoints after erasing alternative feature contrast in 890 reference EXECUTE-to-ASK pairs. Bars show estimates and 95% scene-bootstrap intervals.

Table 1: Matched memory comparison. The two NF models differ only in the historical information they receive.

Memory	ID accuracy (%) ↑	Non-MNL accuracy (%) ↑
Choice context	77.00	75.97
Chosen only	58.07	56.53
Gain	+18.93	+19.44

4.3 ERASING ALTERNATIVE CONTRAST CHANGES AGENT DECISIONS

We next replace every unchosen feature vector in each historical event with the selected vector, while keeping the selected history, current candidates, masks, scene, and event size fixed. This preserves the stored outcome but removes the feature contrast carried by the alternatives.

Among 890 reference EXECUTE-to-ASK transitions, AFFORDANCEMEM-BANK recovers both endpoints in 92.13% of pairs. Laplace and MAP recover 77.64% and 73.03%, respectively (Figure 3b). Chosen-only and no-memory methods remain unchanged because they do not observe the modified alternatives. This intervention links the retained comparison evidence to the later decision.

4.4 FROM PREFERENCE EVIDENCE TO SELECTIVE ACTION

At 70% autonomous coverage, AFFORDANCEMEM-BANK with the regret policy has 8.19% hidden-target error, compared with 15.03% for MAP with entropy (Figure 5a). This compares complete selective strategies, rather than an isolated gating rule.

On the human-study data, with clarification cost 0.10, the selective policy acts on 38.87% of trials and reaches mean cost 0.0862, compared with 0.1000 for always asking and 0.1637 for always acting (Figure 5b). Separately, Table 2 reports the ID hidden-utility trade-off across three clarification costs.

4.5 TRANSFER ACROSS SESSIONS, MECHANISMS, AND SCENES

Repeated interaction. We evaluate 96 human participants over two sessions, 48 hours apart, using Session-A choices to predict their actual Session-B choices. On successfully recorded Session-B choices, choice-context memory reaches 54.88% top-1 accuracy, compared with 47.53% for matched

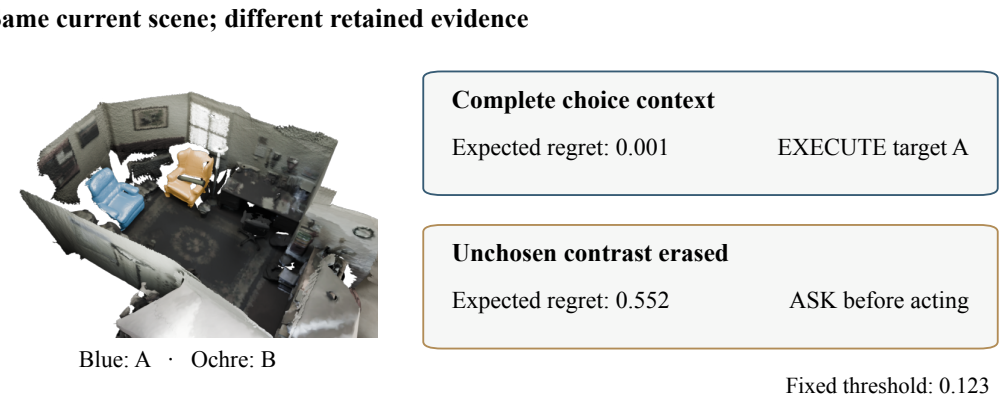

Figure 4: **The same scene can call for acting or asking.** For one saved AFFORDANCEMEM-NF trace, the complete history gives expected regret 0.001 and executes target A. Erasing only the unchosen contrast raises regret to 0.552 and triggers clarification.

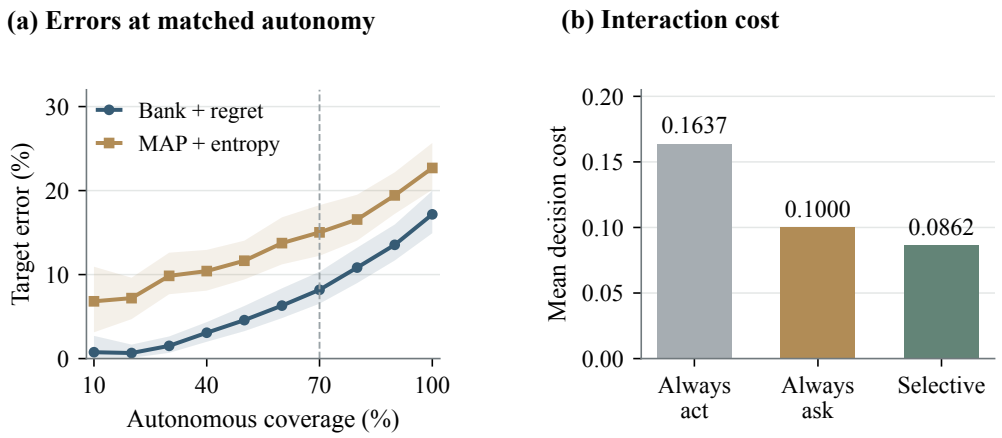

Figure 5: **Preference evidence supports selective autonomy.** (a) Hidden-target error versus autonomous coverage. (b) Mean decision cost on the human-study data with clarification cost 0.10. Shaded regions show 95% scene-bootstrap intervals where applicable.

chosen-only memory. Laplace is slightly higher overall, showing that memory representation and belief estimation are separate design choices.

Choice mechanisms. Across six non-MNL generators, AFFORDANCEMEM-BANK reaches 79.89% hidden-target accuracy with regret 0.2033, compared with 76.65% and 0.2629 for Laplace. On the four preference shifts, Bank is stronger on lexicographic and pairwise preferences, while Laplace is stronger on dense and drifting preferences. No estimator dominates every mechanism; the matched comparisons above provide the direct test of memory content.

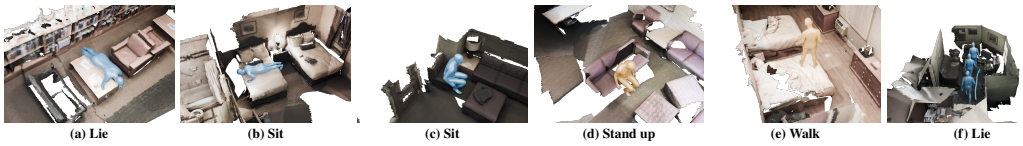

Figure 6: **Rendered scene and motion context.** Six HUMANISE–ScanNet examples show approved lie, sit, stand-up, and walk targets and their supplied motion executions.

Table 2: ID hidden-utility decision cost / autonomous coverage at representative clarification costs.

Cost	AFFORDANCEMEM-BANK	Laplace	MAP
0.05	0.0370 / 46.6%	0.0515 / 6.1%	0.0505 / 10.1%
0.10	0.0632 / 56.7%	0.0950 / 26.3%	0.0965 / 15.3%
0.20	0.0997 / 72.1%	0.1534 / 47.9%	0.1587 / 55.8%

Table 3: Two-session results for 96 human participants, evaluated on successfully recorded Session-B choices.

Method	Top-1 (%) ↑	Acceptable (%) ↑	Regret ↓
Laplace MNL	**55.19**	**82.66**	**0.314**
AFFORDANCEMEM-BANK	54.88	82.50	0.321
MAP MNL	52.22	78.39	0.385
Matched chosen-only	47.53	72.24	0.499
No memory	29.81	54.43	0.918

Scene and candidate construction. LINGO provides language, SMPL-X motion, occupancy scenes, and endpoints (Jiang et al., 2024a). Clustering endpoints into 2–6 goal regions produces 10,920 records from 108 canonical scenes. With HUMANISE thresholds frozen, AFFORDANCEMEM-BANK reaches 74.26% hidden-target accuracy on 607 test sources at a 0.4 m radius, versus 37.89% for chosen-only memory and 24.05% for no memory. Laplace reaches 63.43%; the Bank–Laplace gap remains above nine points at 0.3 and 0.5 m. Appendix C.4 reports action-stratified results.

5 DISCUSSION AND CONCLUSION

Our results support a simple view of personalized agent memory: when a user chooses among alternatives, the comparison itself can provide useful evidence for future decisions. Storing only the selected outcome can erase this evidence and make different choice situations appear identical. Across parameter-matched models, preserving choice context improves later target selection. A targeted intervention shows the same effect more directly: changing only the unchosen alternatives can change whether the agent executes or asks, while the selected history and current task remain fixed.

This shifts the role of memory from storing past outcomes to preserving evidence for later decisions. The preference model is one way to interpret that evidence, and our finite and learned belief models behave differently across mechanisms. The more stable observation is that memory content matters. Choice-context memory is therefore complementary to better retrieval, updating, or organization, which can only use information that survives memory construction.

The act-or-ask setting makes the downstream consequence of memory content explicit. A useful memory should not only improve target prediction, but also help the agent decide whether its evidence is sufficient for autonomous action. Repeated interactions and LINGO transfer suggest that this principle extends beyond one episode construction or choice model. At the same time, no single belief estimator is best across all behaviors. This separates two questions: what evidence should be preserved, and how it should be interpreted.

AffordanceMem focuses on this memory-to-decision interface. Candidate proposal, feasibility checks, and motion execution are supplied components, and our preference evaluations combine controlled benchmarks with a two-session human study. The offered set is the observed comparison context; we do not infer which alternatives the user actually considered. These boundaries isolate the effect of memory content while leaving richer user behavior, learned candidate construction, and end-to-end deployment for future work.

In summary, a choice is more than its outcome. For personalized agents, the alternatives surrounding a choice can determine what the interaction teaches the system and what action it should take next. Choice-context memory preserves this decision-relevant evidence from past interaction to future action.

AI USE STATEMENT

AI tools were used for file retrieval and manuscript language polishing. The authors reviewed the AI-assisted work and take responsibility for the final manuscript.

REPRODUCIBILITY STATEMENT

Section 4 describes the evaluation setting and metrics. Appendix A provides the proof and analytic example; Appendices B and C provide experimental details and additional controls and transfer results.

REFERENCES

David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24 (2):265–272, 1953.

Zhi Cen, Huaijin Pi, Sida Peng, Zehong Shen, Minghui Yang, Shuai Zhu, Hujun Bao, and Xiaowei Zhou. Generating human motion in 3D scenes from text descriptions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1855–1866, 2024.

Xingyu Chen, Rui Wang, Zhaopeng Tu, and Liefeng Bo. AdaMem: Learning what to remember with adaptive memory policies for personalized agents. *arXiv preprint arXiv:2606.21144*, 2026. URL https://arxiv.org/abs/2606.21144.

Jiayang Cheng, Dongyu Ru, Lin Qiu, Yiyang Li, Xuezhi Cao, Yangqiu Song, and Xunliang Cai. AMemGym: Interactive memory benchmarking for assistants in long-horizon conversations. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=sfrVLzsmlf.

Angela Dai, Angel X. Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Niessner. ScanNet: Richly-annotated 3D reconstructions of indoor scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5828–5839, 2017.

Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. Density estimation using real NVP. In *International Conference on Learning Representations*, 2017. URL https://arxiv.org/abs/1605.08803.

Ran El-Yaniv and Yair Wiener. On the foundations of noise-free selective classification. *Journal of Machine Learning Research*, 11:1605–1641, 2010.

Rachel Freedman, Rohin Shah, and Anca D. Dragan. Choice set misspecification in reward inference. In *Proceedings of the Workshop on Artificial Intelligence Safety 2020*, volume 2640 of *CEUR Workshop Proceedings*, 2020. URL https://ceur-ws.org/Vol-2640/paper_14.pdf.

Yonatan Geifman and Ran El-Yaniv. SelectiveNet: A deep neural network with an integrated reject option. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 2151–2159, 2019.

Siyuan Huang, Zan Wang, Puhao Li, Baoxiong Jia, Tengyu Liu, Yixin Zhu, Wei Liang, and Song-Chun Zhu. Diffusion-based generation, optimization, and planning in 3D scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 16750–16761, 2023.

Nan Jiang, Zimo He, Zi Wang, Hongjie Li, Yixin Chen, Siyuan Huang, and Yixin Zhu. Autonomous character–scene interaction synthesis from text instruction. In *SIGGRAPH Asia 2024 Conference Papers*, pp. 1–11, 2024a.

Nan Jiang, Zhiyuan Zhang, Hongjie Li, Xiaoxuan Ma, Zan Wang, Yixin Chen, Tengyu Liu, Yixin Zhu, and Siyuan Huang. Scaling up dynamic human–scene interaction modeling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1737–1747, 2024b.

Ananth Jonnavittula and Dylan P. Losey. I know what you meant: Learning human objectives by (under)estimating their choice set. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2747–2753, 2021. doi: 10.1109/ICRA48506.2021.9562048.

Jiho Kim, Junseong Choi, Woosog Chay, Daeun Kyung, Yeonsu Kwon, Yohan Jo, and Edward Choi. ProPerSim: Developing proactive and personalized AI assistants through user-assistant simulation. In *International Conference on Learning Representations*, 2026.

Belinda Li, Alex Tamkin, Noah Goodman, and Jacob Andreas. Eliciting human preferences with language models. In *International Conference on Learning Representations*, 2025.

Stella Li, Avinandan Bose, Faeze Brahman, Simon Du, Pang Wei Koh, Maryam Fazel, and Yulia Tsvetkov. PrefDisco: Benchmarking proactive personalized reasoning. In *International Conference on Learning Representations*, 2026.

R. Duncan Luce. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, 1959.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. In Paul Zarembka (ed.), *Frontiers in Econometrics*, pp. 105–142. Academic Press, 1974.

Sahand Negahban, Sewoong Oh, Kiran K. Thekumparampil, and Jiaming Xu. Learning from comparisons and choices. *Journal of Machine Learning Research*, 19(40):1–95, 2018.

Zhuoshi Pan, Qianhui Wu, Huiqiang Jiang, Xufang Luo, Hao Cheng, Dongsheng Li, Yuqing Yang, Chin-Yew Lin, H. Vicky Zhao, Lili Qiu, and Jianfeng Gao. SeCom: On memory construction and retrieval for personalized conversational agents. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/e56f394bbd4f0ec81393d767caa5a31b-Abstract-Conference.html.

Allen Z. Ren, Anushri Dixit, Alexandra Bodrova, Sumeet Singh, Stephen Tu, Noah Brown, Peng Xu, Leila Takayama, Fei Xia, Jake Varley, Zhenjia Xu, Dorsa Sadigh, Andy Zeng, and Anirudha Majumdar. Robots that ask for help: Uncertainty alignment for large language model planners. In *Proceedings of the 7th Conference on Robot Learning*, pp. 661–682, 2023.

Arjun Seshadri, Alex Peysakhovich, and Johan Ugander. Discovering context effects from raw choice data. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 5660–5669, 2019. URL https://proceedings.mlr.press/v97/seshadri19a.html.

Rajeev Verma and Eric Nalisnick. Calibrated learning to defer with one-vs-all classifiers. In *Proceedings of the 39th International Conference on Machine Learning*, pp. 22184–22202, 2022.

Yuan Wang, Yali Li, Xiang Li, and Shengjin Wang. HSI-GPT: A general-purpose large scene–motion–language model for human scene interaction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7147–7157, 2025.

Zan Wang, Yixin Chen, Tengyu Liu, Yixin Zhu, Wei Liang, and Siyuan Huang. HUMANISE: Language-conditioned human motion generation in 3D scenes. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Nils Wilde, Dana Kulic, and Stephen L. Smith. Active preference learning using maximum regret. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 10952–10959, 2020. doi: 10.1109/IROS45743.2020.9341530.

Di Wu, Hongwei Wang, Wenhao Yu, Yuwei Zhang, Kai-Wei Chang, and Dong Yu. LongMemEval: Benchmarking chat assistants on long-term interactive memory. In *International Conference on Learning Representations*, 2025. URL https://arxiv.org/abs/2410.10813.

Wujiang Xu, Zujie Liang, Kai Mei, Hang Gao, Juntao Tan, and Yongfeng Zhang. A-MEM: Agentic memory for LLM agents. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-0593.

Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Ruslan Salakhutdinov, and Alexander J. Smola. Deep sets. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Shuhan Zhang, Wendi Ren, and Shuang Li. Logic-Logit: A logic-based approach to choice modeling. In *International Conference on Learning Representations*, 2025.

Kaifeng Zhao, Shaofei Wang, Yan Zhang, Thabo Beeler, and Siyu Tang. Compositional human–scene interaction synthesis with semantic control. In *European Conference on Computer Vision*, pp. 311–327, 2022.

Siyan Zhao, Mingyi Hong, Yang Liu, Devamanyu Hazarika, and Kaixiang Lin. Do LLMs recognize your preferences? evaluating personalized preference following in LLMs. In *International Conference on Learning Representations*, 2025.